technical note technical note techn

Effect of Free Flight Conditions on Controller Performance, Workload, and Situation Awareness

Mica R. Endsley, Ph.D., Texas Tech University
Richard H. Mogford, ACT-530
Kenneth R. Allendoerfer and
Michael D. Snyder, NYMA, Inc.
Earl S. Stein. Ph.D., ACT-530

U.S. Department of Transportation
Federal Aviation Administration

William J. Hughes Technical Center
Atlantic City International Airport, NJ 08405

NOTICE

This document is disseminated under the sponsorship
of the U.S. Department of Transportation in the interest of
information exchange. The United States Government
assumes no liability for the contents or use thereof.

The United States Government does not endorse
products or manufacturers. Trade or manufacturers'
names appear herein solely because they are considered
essential to the objective of this report.

1. Report No. DOT/FAA/CT-TN97/12	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Effect of Free Flight Conditions on Controller Performance, Workload, and Situation Awareness		5. Report Date December 1997	
		6. Performing Organization Code ACT-530	
7. Author(s) Mica R. Endsley, Texas Tech University, Richard H. Mogford, ACT-530, Kenneth R. Allendoerfer and Michael D. Snyder, NYMA, Inc., and Earl Stein, ACT-530		8. Performing Organization Report No. DOT/FAA/CT-TN97/12	
9. Performing Organization Name and Address Federal Aviation Administration William J. Hughes Technical Center Atlantic City International Airport, NJ 08405		10. Work Unit No. (TRAIS)	
		11. Contract or Grant No. F2202J	
12. Sponsoring Agency Name and Address Federal Aviation Administration Office of Aviation Research, Human Factors Division 800 Independence Ave., S.W. Washington, DC 20591		13. Type of Report and Period Covered Technical Note	
		14. Sponsoring Agency Code AAR-100	
15. Supplementary Notes			

16. Abstract

Free flight represents a major change in the way that aircraft are handled in the National Airspace System. It has the potential to significantly increase airspace utilization and, by doing so, improve aircraft throughput. The degree to which these objectives can be met without compromising aircraft safety will depend on appropriate changes in the air traffic control system. This study provides an evaluation of some of the potential effects of free flight on controllers' ability to maintain an accurate and complete picture of the traffic situation. This picture or mental representation is essential for monitoring and separation functions. The study revealed that, using current technology, some aspects of free flight may adversely influence the situation awareness and performance of controllers. The results provide information on some possible consequences of free flight that should be explored in future research.

17. Key Words Free Flight Situation Awareness Air Traffic Control		18. Distribution Statement This document is available to the public through the National Technical Information Service, Springfield, Virginia, 22161	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 53	22. Price

Form DOT F 1700.7 (8-72)   **Reproduction of completed page authorized**

Table of Contents

iii

Appendixes

A - Air Traffic Control Evaluation: SME Observation Form
B - Situation Awareness Global Assessment Technique Queries
C - NASA-Task Load Index
D - Post-Scenario Questionnaire
E - SME SAGAT Data Evaluation Form
F - Exit Questionnaire

List of Illustrations

Figures
Page

Executive Summary

Free flight represents a major change in the way that aircraft are handled in the National Airspace System. It has the potential to significantly increase airspace utilization and, by doing so, improve aircraft throughput. The degree to which these objectives can be met without compromising aircraft safety will depend on appropriate changes in the air traffic control (ATC) system. This study provides an evaluation of some of the potential effects of free flight on controllers' ability to maintain an accurate and complete picture of the traffic situation. This mental representation is essential for monitoring and separation functions.

Facilities at the Research Development and Human Factors Laboratory (RDHFL) at the William J. Hughes Technical Center were used for the study. ATC simulation scenarios were created reflecting present-day operations and three levels of free flight. A sector from Jacksonville Air Route Traffic Control Center (ARTCC) airspace was used. In the baseline (B) condition, traffic followed customary flight paths and procedures. The first level of free flight (D) placed aircraft on direct routes. Deviations around weather and for other purposes were permitted, with clearances from ATC. In the second level of free flight, deviations were taken by the aircraft, which were only required to notify ATC of their intent (D/I). The highest level of free flight did not require the aircraft to notify ATC; they could deviate at will without announcing intent (D/NI).

Dependent variables included performance, situation awareness (SA), workload, and subjective opinion data. Ten controllers from Jacksonville ARTCC participated in the 5-week study with two controllers present per week. Upon arrival at the RDHFL, they were trained on the simulation platform, ATCoach. ATC data were displayed on a 20 X 20 in. monitor and flight strips and bays were provided.

The experimental design was within-subjects with participants working independently in pairs. A training trial for each of the four experimental conditions was administered followed by two data collection trials. Order of presentation was counterbalanced and trials were given over a 3-day period. Workload data were collected every 5 min using the Air Traffic Workload Input Technique (ATWIT). Each scenario was frozen at four randomly selected times to gather SA data using the Situation Awareness Global Assessment Inventory (SAGAT). An over-the-shoulder performance rating form was completed by observers (supervisors from Jacksonville ARTCC) and questionnaires were filled out at the end of each run. Simulation pilots used scripts to inject aircraft deviations in each scenario, according to the experimental condition.

Objective performance results indicated a total of seven operational errors in the D/NI condition, as compared to six in the other three conditions combined. Due to the small number of operational errors, which is typical in ATC simulations, no statistical conclusions could be drawn. However, this result is worthy of concern.

Although controllers indicated no differences in their performance across conditions, observers rated performance significantly lower in the D/I and D/NI free flight conditions, as compared to the B condition. Controllers were observed to perform less well in marking flight strips and prioritizing tasks in the D/NI condition. Controllers were rated as providing significantly less

information in the D/I and D/NI conditions, as compared to the B conditions. They also made more speed changes in the B condition as compared to the three free flight conditions. Other objective and subjective measures of efficiency, safety, and traffic flow showed no difference between conditions.

Free flight had an impact on controller workload. The National Aeronautics and Space Administration Task Load Index scores and a posttest measure showed that workload was rated higher by participants at higher levels of free flight. There was no effect on ATWIT scores. Observer ratings showed better performance in detecting pilot deviations and overall attention and SA in the B condition as compared to the D/NI condition.

Analyses of variance conducted on the SAGAT data revealed significant losses in SA on some variables at higher levels of free flight. These included knowledge of aircraft location, callsign, next sector for aircraft, weather impact, aircraft with uncompleted clearances, correct receipt of aircraft clearances, and conformance to clearances. Controllers were aware of fewer aircraft, and, for those aircraft, they displayed lower SA particularly at the higher levels of SA (comprehension and projection).

This study showed that, if controllers are expected to act as passive monitors of free flight air traffic, their awareness of the state of air traffic may be reduced, their workload may increase, and their ability to intervene in a timely manner may be somewhat limited. The results reported here are probably indicative of a lower degree of predictability or SA associated with the free flight concepts evaluated. The use of technologies and displays for dealing with these concerns should be explored. It is also possible that compensating mechanisms developed through practice and experience may be found to provide the levels of SA needed for adequate functioning under free flight.

This study is one of the first conducted to look at free flight systematically. It may reveal the impact of such concepts on the ability of controllers and pilots to function in a safe and efficient manner. More such studies are needed to expand this effort and examine other aspects of free flight and its consequences on the air transportation system.

1. Introduction

With the advent of new technologies such as the Global Positioning System and a cockpit display of traffic information, the concept of free flight may be introduced as a major change in the way that air traffic is managed in United States air space. The RTCA provides one idea of how free flight might be implemented (RTCA, 1995). Under free flight, aircraft would no longer be restricted to flying the air corridors, which comprise only approximately 5% of available airspace. Aircraft pilots would have more control over setting their routes and over making dynamic changes in the flight path, altitude, and speed of their aircraft while under instrument flight rule conditions. The ability to fly to destinations directly instead of along fixed routes may create significant advantages for both time and fuel savings for the operators of aircraft and, particularly, for major airlines. It also allows pilots to have more control in avoiding weather and dealing with other factors that may emerge during a flight.

The exact way in which free flight will be implemented has not yet been determined. Opinions on the changes in procedures, displays, and automation needed to support free flight vary as do the concepts regarding the new roles of the pilot and the controller under free flight. Information on the degree to which a given concept can be accomplished without compromising aircraft safety should drive any decision on whether and how to implement free flight. Airspace planners must consider many factors in making this decision.

1.1 Background

There are significant challenges to providing aircraft with sufficient information to be able to make decisions that do not place them into an unsafe proximity with other aircraft. In addition, controller roles under free flight will be different. As their ability to control the actions and paths of aircraft evolves, the controllers' role is anticipated to change to that of monitor, taking action only when separation problems are detected (RTCA, 1995). The ability of the controller to perform under these condition needs to be evaluated as the controller's situation awareness (SA) may significantly change under free flight conditions.

The concept of free flight represents a change in the dynamics and behavior of the aircraft operating in a controller's sector. With free flight, it is likely that the ability of the controller to determine why an aircraft is behaving in a particular way will be reduced. For example, does a deviation of an aircraft from its current path represent an intentional action or a problem of which the pilot is not aware? Is the pilot aware of potential conflicts or altitude problems? Not only will controllers need to be able to detect changes in aircraft flight path, speed, and altitude, they will need to assess the impact on separation with other aircraft, special airspace, or given standards (e.g., airport approach volume limits or minimum altitude restrictions). Being able to do so in the current system depends on understanding aircraft intent and pilot expectation.

Acquiring this kind of information under free flight conditions may increase communications requirements and alter the behavior of the controller. Communications may occur more for the purpose of information exchange with the pilot instead of issuing control motivations. These transactions may be much more frequent and time consuming. An increase in controller

workload may result from this change as opposed to the decrease in workload that is generally assumed to occur under free flight. There will likely be significant new demands for the controller to be able to interpret aircraft actions and understand their significance for aircraft safety. This may create a real possibility that SA will be degraded if the controller cannot keep up with these demands.

In addition, the predictability of aircraft movement will likely decrease under free flight. In the current system, controllers gain information about how the aircraft is going to behave from knowledge of their assigned flight path and destination. There are a limited number of ways that aircraft will proceed through a given airspace according to a given flight path and the aircraft intended activity in that sector (e.g., approach, departure, or en route). The controller can usually detect deviations from these norms quickly. With free flight, aircraft may come from almost any direction into a sector, change paths many times without controller action or approval, and depart the sector in almost any direction. With this loss of aircraft predictability, the ability of the controller to determine potential separation problems may be reduced. Projection of the future actions of aircraft (the highest level of SA) is critical to the controller's ability to make timely control actions. Thus, there is a significant concern as to controllers' ability to understand the significance of aircraft actions and adequately predict impending problems, allowing them to manage traffic effectively.

Currently, it is unclear what will be the spectrum of enabling technologies. It is also not known what compensating mechanisms controllers may bring to bear on these types of problems. Compensation may occur in the form of new procedures for facilitating the greater flow of information this system will necessitate. Controllers may also adopt new strategies for controlling traffic under these types of conditions. In order to determine what compensation might be appropriate, however, concrete information on the actual effects of free flight is needed. To date, most information on free flight is highly unstructured and analytic.

1.2 Objective and Scope

The objective of this study was to examine the effect of free flight concepts on controllers' ability to create and maintain an accurate picture of the air traffic situation. We also examined its impact on controller workload, control strategies, and performance. The study was accomplished using current technology and controller capabilities. This should be viewed as an initial investigation of certain free flight concepts and not as an assessment of free flight in its entirety.

New air traffic control (ATC) or cockpit technologies should be evaluated separately from the operational concept changes that were the focus here. As no new cockpit technologies were specifically investigated, this study did not examine the ability of pilots to separate themselves from other traffic nor free flight feasibility from this standpoint. The objective of this study was restricted to examining the ability of the controller to maintain SA and provide traffic separation while working in a hypothetical free flight environment.

2. Method

Free flight encompasses many new operational concepts and new technologies in a constantly evolving manner. Therefore, only a few critical aspects were selected for this study. These included the use of direct routes, the ability of pilots to deviate from flight plans of their own accord, and the requirement to inform controllers of pilot intentions in making such deviations. The experiment was designed so that the impact of each factor could be isolated from the others. Four conditions representing increasingly higher levels of free flight were examined.

2.1 Experimental Design

A within-participants design was used with four levels of one independent variable. Conditions were administered in a semi-counterbalanced order.

2.1.1 Independent Variable

The level of free flight provided served as the independent variable for the study. All other aspects of the ATC system, including procedures and technologies, were held as consistent as possible between conditions. The level of free flight was based on four conditions:

- Baseline (B) - This condition employed current ATC procedures for controlling traffic. This involved all normal procedures active in the modeled sector at the time of the study. Pilots filed normal route plans and could only deviate from those plans with a clearance from the controller (although they could request deviations, which the controller could choose to grant or deny, as appropriate).

- Direct Routing (D) - This condition incorporated similar scenarios (same traffic density and complexity) as the B condition, but aircraft were provided with flight plans with direct routings. (Slight modifications to some direct routes were made, if needed, so that no flight plans were filed through restricted areas).

- Direct Routings/Deviations with Intent (D/I) - This condition incorporated scenarios similar to the Direct Routing condition, however, pilots were also allowed to deviate from their filed routes at will after conveying their intentions to the controller (through verbal radio transmission). The controller's role was to reject or modify such deviations only if necessary to insure safety of flight (i.e., on an exception basis).

- Direct Routings/Deviations without Intent (D/NI) - This condition incorporated scenarios similar to the Direct Routings/Deviation with Intent condition. Pilots were not required to convey their intentions to the controller in advance of making deviations from their filed flight path but could simply deviate at will. The controller's role was to reject or modify such deviations when detected only if necessary to insure safety of flight (i.e., on an exception basis).

2.1.2 Dependent Variables

Several measures were examined as dependent variables, including performance, SA, workload, and the subjective impressions of the controllers serving as participants in this study.

2.1.2.1 Performance

Objective performance data were collected by the simulation computer during the study, and calculations were performed to derive performance measures. They were

- a. Safety of Flight
    1. number of operational errors.

- b. Efficiency
    1. number of flights handled,
    2. duration of flights handled,
    3. distance flown in sector,
    4. number of completed flights,
    5. number of aircraft holds,
    6. duration of aircraft holds,
    7. number of successful hand-offs, and
    8. number of hand-off misses.

- c. Control Strategy
    1. number changes in altitude/aircraft handled,
    2. number changes in speed /aircraft handled, and
    3. number changes in heading/aircraft handled.

- d. Taskload
    1. number controller entries,
    2. number controller transmissions, and
    3. duration of controller transmissions.

A subjective measure of performance was obtained through a rating of each participant on the Observation Form at the conclusion of each trial (see Appendix A). The Observation Form was developed by the FAA Research Development and Human Factors Laboratory (RDHFL) at the Federal Aviation Administration (FAA) William J. Hughes Technical Center. It is used by subject matter experts (SMEs) to make over-the-shoulder evaluations of controller performance during ATC simulations (Sollenberger, Stein, & Gromelski, 1997). Subjective ratings of the participants' performance were made on an eight-point scale (1- extremely poor judgment and made very

frequent errors to 8 - always demonstrated excellent judgment and used outstanding control techniques). The factors were

    a. Maintaining Safe and Efficient Traffic Flow

        1. maintaining separation and resolving potential conflicts,

        2. sequencing arrival and departure aircraft efficiently,

        3. using control instructions effectively, and

        4. overall.

    b. Maintaining Attention and SA

        1. maintaining awareness of aircraft positions,

        2. ensuring positive control,

        3. detecting pilot deviations from control instructions,

        4. correcting own errors in a timely manner, and

        5. overall.

    c. Prioritizing

        1. taking actions in an appropriate order of importance,

        2. preplanning control actions,

        3. handling control tasks for several aircraft,

        4. marking flight strips while performing other tasks, and

        5. overall.

    d. Providing Control Information

        1. providing essential air traffic control information,

        2. providing additional air traffic control information, and

        3. overall.

    e. Technical Knowledge

        1. showing knowledge of LOAs and SOPs,

        2. showing knowledge of aircraft capabilities and limitations, and

        3. overall.

 f. Communicating

  1. using proper phraseology,

  2. communicating clearly and efficiently,

  3. listening for pilot readbacks and requests, and

  4. overall.

In addition, the SMEs provided a rating describing the participant's control strategy on a 10- point scale (1-none to 10-always). The various strategies were

 a. Preference for Vertical Separation,

 b. Preference for Separation Through Vectoring,

 c. Preference for Speed Control, And

 d. How Well the Controller Controlled Traffic (1-poor to 10- extremely well).

The participants also subjectively rated their own performance on the same 10-point scale (1-poor to 10- extremely well) and provided subjective comments on the realism of the simulation and free flight conditions.

2.1.2.2 Situation Awareness

The Situation Awareness Global Assessment Technique (SAGAT): ATC Version (Endsley & Kiris, 1995a) was used to measure participant SA during the test. Four randomly placed freezes were inserted into each trial to collect SAGAT data. Each of the following SAGAT queries were administered at each stop (Appendix B). These queries consisted of

 a. Level 1 SA - Perception of the Traffic Situation

  1. aircraft location,

  2. aircraft level of control,

  3. aircraft callsign,

  4. aircraft altitude,

  5. aircraft groundspeed,

  6. aircraft heading,

  7. aircraft flight path change (vertical, turning), and

  8. aircraft type.

 b. Level 2 & 3 SA - Comprehension & Projection of Traffic Situation

  1. aircraft next sector,

  2. aircraft separation,

  3. aircraft assignments,

4. assignment reception,

5. aircraft conformance,

6. aircraft hand-offs,

7. aircraft communications,

8. special airspace separation, and

9. weather impact.

2.1.2.3 Workload

The Air Traffic Workload Input Technique (ATWIT) (Stein, 1985) was administered at 5-min intervals throughout the trial to obtain a subjective workload rating from the participants on a 10-point scale. The participant's subjective experienced workload was also assessed immediately following each trial using the National Aeronautics and Space Administration (NASA) Task Load Index (NASA-TLX) (Hart & Staveland, 1988) (see Appendix C). In addition, a subjective assessment of workload was made at the end of each trial on a 10-point scale by both the SMEs and the participants on the subjective questionnaires.

2.1.2.4 Subjective Questionnaire

The Post-Scenario Questionnaire was provided to participants after completing each trial to obtain their evaluation of the ease of controlling traffic, level of awareness of the traffic situation, and perceived performance level in the preceding trial (Appendix D). Information on the viability of each free flight condition tested and recommendations for needed modifications were elicited at the completion of each test condition.

2.2 Participants

The participants included 10 Full Performance Level Air Traffic Control Specialists from the Air Route Traffic Control Center (ARTCC) located in Jacksonville (ZJX). All participants were current at the center with at least 16 hours of operational time in the month preceding testing and with self-reported corrected vision of at least 20/30. Participation in the study was voluntary.

2.3 Apparatus

The study was conducted at the FAA William J. Hughes Technical Center RDHFL using the ATCoach Version 7.0 (1996) simulation system operating on a Sun workstation. ATCoach provides a realistic, high-fidelity simulation of a controller's workstation. The ATCoach system includes a 20 by 20 inch high-resolution color radar display monitor (2000 by 2000 pixels), a three-button trackball, and keyboard. A flight strip bay with printed, standard configuration flight strips was provided for each aircraft in the simulation. A touch panel for the ATWIT measure was included to the left of the radar screen. SAGAT and NASA-TLX data were collected via programs running on Hypercard on a Macintosh computer placed adjacent to the controller's station. Subjective measures were gathered using paper forms.

2.4 Simulated Airspace

The simulated airspace used in this study was the Greencove/Keystone combined sector of ZJX. The sector definition and the original traffic scenarios were developed for an earlier study examining generic airspace (Stein & Guttman, in preparation).

At ZJX, the Greencove/Keystone combined sector is responsible for altitudes 24,000 ft and above and has four major traffic flows. Southbound aircraft enter the sector from the northeast and northwest and continue south and southeast toward the cities of Miami, Fort Lauderdale, and West Palm Beach along the J45 or J79 airways. These aircraft are generally at their final altitude when they reach the sector and do not require altitude clearances from the controller. Northbound aircraft leave the Orlando International Airport and head northwest or north along the J81 or J53 airways and generally contact Greencove/Keystone at about 18,000 ft while climbing to an interim altitude of 23,000 ft. The controller working Greencove/Keystone will clear the aircraft to climb to its final altitude when available. Other northbound aircraft depart from southeast Florida and enter the sector in the south, near Orlando. These aircraft continue north and northwest along the J81 and J53 airways. In the field, these aircraft are usually at their final altitude when they reach Greencove/Keystone but occasionally may require altitude clearances from the controller. In this simulation, these northbound aircraft were at their final altitudes when they reached Greencove/Keystone.

Greencove/Keystone is bordered beneath by the St. Johns and St. Augustine sectors, on the northeast by the combined States/Hunter sector, on the north-northwest by the combined Alma/Moultrie sector, on the west by the Lake City/Ocala sector, on the southwest by the Mayo sector; on the south by the Boyel sector of Miami ARTCC (ZMA), and on the south-southeast by the Hobee sector of ZMA. For the purpose of the simulation, all adjacent sectors accepted all hand-offs and approved all point-outs. The sector is bordered on the east by a warning area controlled by the US Navy. In the field, civilian aircraft may enter the warning area only with special permission. For the purpose of this simulation, the warning area was considered active and no civilian aircraft were permitted to enter.

2.5 Traffic scenarios

Fourteen scenarios of simulated air traffic were developed for the sector.

2.5.1 Baseline scenarios

Five of the fourteen scenarios were baseline scenarios using current traffic patterns. In these scenarios, aircraft had standard flight plans containing their departure and arrival airports and all the fixes and airways along their routes of flight. Four of the baseline scenarios were used in the earlier study (Stein & Guttman, in preparation) and the fifth was developed by changing the callsigns of the second baseline scenario. During the experiment, controllers worked four baseline scenarios: one simulator familiarization trial, one B condition practice trial, and two B condition test trials.

2.5.2 Free Flight Scenarios

Nine free flight scenarios were developed using the four original baseline scenarios. To create these scenarios, the flight plans from the four original scenarios were edited to eliminate the intervening fixes between the origin and destination airports. This technique, however, required several modifications and refinements to preserve the traffic volume and simulation fidelity. First, for some flights, the straight line between the airports did not cross through Greencove/Keystone. This was especially true for arrival aircraft from the northwest flying toward southeast Florida. This required changing the arrival or departure airports for these flights so that the aircraft would continue to fly through the sector. For example, a direct flight from Nashville to Miami misses the sector along the extreme southwest corner. This was changed to a flight between Indianapolis and Fort Lauderdale. The general pattern and direction of traffic and the correct altitude for direction of flight were always preserved. Second, aircraft departing the northeast U.S. and flying direct toward southeast Florida would fly much of their course over the Atlantic Ocean and through the various U.S. military warning areas along the U.S. east coast. In order to avoid this, arrival aircraft from the northeast entered Greencove/Keystone at the extreme northeast corner of the sector and proceeded south as close to the warning area as allowed. Though these flights were not technically direct routed, they were as close to direct as possible without eliminating flights from the northeast, requiring long flights over the ocean, or altering the military airspace.

The second step in creating free flight scenarios was to change the aircraft callsigns in the first four scenarios to produce five new scenarios with the same flight plans but different callsigns. The number of major and minor carriers and the frequencies with which particular airlines fly through the Greencove/Keystone sector were preserved.

2.5.3 Deviation Scripts

Each scenario contained two heavy storm cells, which were displayed on the controller's radar and moved during the scenario. The exact shape and speed of the weather system differed in each scenario, but all were judged by the SMEs to be realistic and typical of those encountered at ZJX. Each scenario also contained an area of moderate to heavy turbulence. To avoid the poor weather and ride conditions and improve fuel efficiency, pseudopilots made deviations from their filed flight plans by changing aircraft headings or altitudes. Depending upon the experimental condition, pseudopilots requested controller approval before beginning a deviation (D), informed the controller of their intentions as they began a deviation (D/I), or simply began a deviation without contacting the controller at all (D/NI). Pseudopilots made deviations by following scripts that listed the aircraft to deviate, the approximate simulation time for the deviation, the nature and magnitude of the deviation, and a reason for the deviation. Controllers serving as participants in the study did not have access to these scripts.

Deviation scripts were developed through collaboration among the researchers, the SMEs, and the pseudopilots. To develop pseudopilot deviation scripts, SMEs controlled traffic in each scenario, recorded when deviation requests typically would be made, and noted the nature of the request. This information was adapted into the pseudopilot deviation scripts. Each script contained from 13 to 18 deviations, depending upon the traffic pattern in the scenario, the location of the storm cells, and the location of the turbulence. Generally, heading changes were

scripted to avoid storm cells and altitude changes were added to avoid turbulence. A small number of deviations used an altitude change to improve fuel efficiency. Each script was thoroughly tested during shakedown sessions with SMEs.

In the B and DR conditions, the scripts listed requests to be made by the pseudopilots. The phraseology of these requests followed current ATC practices. Before a deviation could begin, pseudopilots contacted the controller and requested the deviation, giving the reason for the request. When asking for a heading change, pseudopilots requested a specific heading that would allow them to avoid the storm cells. When requesting an altitude change to avoid turbulence, pseudopilots requested "smooth air." This request allowed the controller to determine what altitude to give and is the terminology used in the current system. When asking for an altitude change for fuel efficiency, pseudopilots requested a specific higher altitude that was appropriate for direction of flight. The controller would approve, deny, or change the request. The pseudopilots would then issue the appropriate ATCoach command that would make the simulated aircraft climb, descend, or turn.

For the DI condition, the scripts were modified so that the pseudopilots informed the controller of their intentions and simultaneously issued the appropriate ATCoach command. If the controller called back and denied the action, pseudopilots issued a second ATCoach command to return the aircraft to its original heading or altitude. For example, a pseudopilot might inform the controller that he was climbing 2,000 ft to avoid turbulence. As the call was being made, the pseudopilot entered the command into ATCoach. The controller, aware that there was traffic 2000 ft above, might immediately call back and deny the deviation. The pseudopilot would then issue an ATCoach command to return the aircraft to the original altitude.

For the D/NI condition, scripts were modified further so that the pseudopilots issued ATCoach commands and did not inform the controller of their intentions until the controller specifically asked. Once a controller inquired about a pilot's intentions, the pseudopilots informed the controller of the reason for the deviation. If the controller denied or altered the deviation, the pseudopilots would then enter a new ATCoach command. For example, at 32 min into the run, a pseudopilot might issue an ATCoach command to climb to 39,000 ft. As the aircraft began its climb, the controller would notice the climb and might immediately call the pseudopilot and ask him to "say intentions." The pseudopilot might explain that he was climbing to 39,000 ft for improved fuel efficiency. The controller could then instruct the pseudopilot to return to his original altitude, provide an alternative altitude, or allow the deviation to continue.

During the experimental runs, pseudopilots were instructed to follow the scripts as closely as possible but, if needed, could adapt the scripted deviations to fit the dynamic situation. For example, the script might call for USA123 to request a 10-degree turn to avoid a heavy storm cell at 15 min into the run. However, the controller might have issued a heading change to USA123 earlier in the run, making a 10-degree turn at 15 min too extreme. Pseudopilots, then, would modify the magnitude or timing of the deviation so that the desired effect was still obtained (i.e., the aircraft avoided the storm cell).

2.5.4 Scenario Validation

All scenarios, especially the free flight scenarios developed for this study, were subject to an extensive testing and shakedown period. Six SMEs, three from ZJX and three from other facilities, were involved in the shakedown and worked each scenario several times until no significant errors or inconsistencies remained.

2.6 Procedures

Each participant controlled traffic in all four conditions, which were administered in a semi-counterbalanced order. Following a training trial, two trials for each condition were administered consecutively (each using a different scenario). The five scenarios for the B condition and the nine scenarios for the free flight conditions were assigned to the training or the two test trials randomly for each participant.

The study was conducted over 3 consecutive days for each subject. A lunch break and two to three rest breaks were provided each day. On the initial day of the study, participants were given an introduction to the simulator and study instructions. They were also provided with instructions for SAGAT- and NASA-TLX-paired comparison rating forms. They then had 1 hour of familiarization with the ATCoach simulator. For each condition, participants received training on the free flight condition being tested. This consisted of instructions followed by a 1-hour practice period with two stops to practice filling out SAGAT. Training was immediately followed by the two trials for that condition.

The participant was prompted by an audio tone at 5-minute intervals during each trial to make an entry on the ATWIT scale. Four freezes were placed in each scenario at random times to collect SAGAT data. At the time of each freeze, the radar screen was blanked and the simulation was paused while the participant completed the SAGAT queries. Participants first were provided with a map of the sector, which showed only boundaries and navigation fix points. They were asked to indicate where all aircraft in the sector were on the map (for all aircraft currently under their control in their sector boundaries, recently handed-off aircraft, and for aircraft soon to be in their control). The remaining queries were then asked in random order in relation to the aircraft the participants indicated were present (see Appendix B). Subjects completed all queries and then returned to the simulation at the point where they had left off. They were given a few seconds to observe the radar screen prior to resuming the simulation. At the same time that participants filled out the SAGAT battery, the SME filled out a SAGAT data collection form (while viewing the frozen radar screen and flight strips) to supplement the data collected by the simulation computer (see Appendix E).

At the end of each trial, the SME filled out the Observation Form, and the participant completed the NASA-TLX form. At the end of the two test trials for each condition, the Post-Scenario Questionnaire was also completed by the participant. An Exit Questionnaire was provided to each participant at the conclusion of the experiment (Appendix F).

11

3. Results

The results from the study will be presented in four main sections: ATC performance, controller workload, controller SA, and evaluations of the quality of the simulation and scenarios investigated. Statistical analyses of the data were generally conducted using two-way (condition by trial) repeated-measures analyses of variance (ANOVAs). (Exceptions are noted below.) F statistics are reported if a significant main effect of condition was found. Post hoc comparisons between condition means were conducted using the Tukey-HSD procedure. An alpha level of .05 was used for all statistical tests.

3.1 Air Traffic Control Performance Results

3.1.1 Safety of Flight

3.1.1.1 Objective Data: Safety

Thirteen Operational Errors (OEs)[1] occurred during the 80 experimental runs. The distribution of these conflicts across the experimental conditions is shown in Figure 1. More OEs occurred in the D/NI condition than in the other three conditions, B, D, and D/I, combined.

The Cochran Q test (a non-parametric test designed for use with related samples and small sample sizes) was used to determine if the distribution of OEs differed from chance. To compute the Cochran Q statistic for related samples, each simulation run was given a value of 0 or 1 to indicate whether or not an OE occurred for that participant in that condition. The Cochran Q test failed to show a statistically significant difference between the four conditions, $Q(3, N = 10) = 6.75$, $p = .080$. It does, however, present a trend nearing the established level of significance. Due to the small number of participants and the small number of OEs, this result should be viewed with caution. With a larger number of participants, this distribution may have yielded a statistically significant difference between conditions. Participants appeared to have greater difficulty in maintaining separation under the D/NI condition.

3.1.1.2 Subjective Data: Safety

The SMEs rated how safely and efficiently the participants controlled traffic. No significant difference was found between the four conditions on the two questions pertaining to safety of flight: maintaining separation and resolving potential conflicts and overall safe and efficient traffic flow.

---

[1] An OE is defined as less than 2,000 ft vertical and less than 5 mi horizontal separation above 29,000 ft or less than 1,000 ft vertical and less than 5 mi horizontal separation below 29,000 ft.
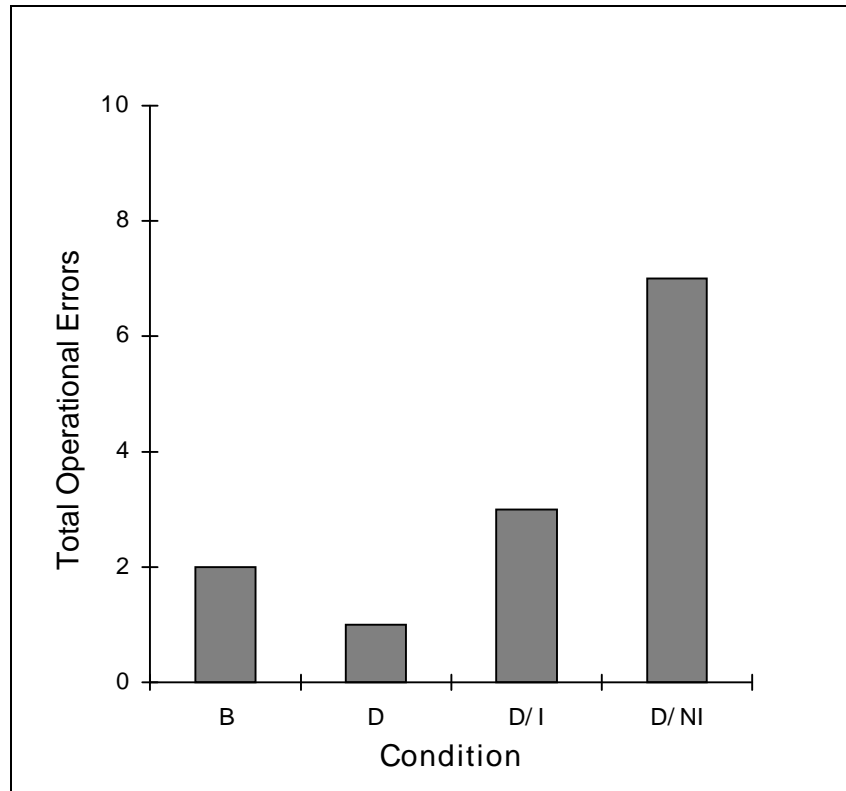
Figure 1. Occurrence of operational errors across conditions.

3.1.2 Efficiency and Prioritizing

3.1.2.1 Objective Data: Efficiency

No significant difference was found between the four conditions on any of the objective performance variables related to efficiency: number of flights handled, duration of flights handled, total distance flown, number of completed flights, number of holds, duration of holds, number of successful hand-offs, and number of hand-off misses.

3.1.2.2 Subjective Data: Efficiency

SMEs rated how effectively participants used control instructions on the using control instructions effectively item on the Observation Form. They also rated the participants' performance in sequencing arrival and departure aircraft efficiently. No significant differences were found between the four conditions on these items.

Although this study did not fully examine many efficiency issues that might occur as a result of free flight, no overall efficiency gains were observed as a result of the provision of direct routings or from allowing pilots to deviate from flight plans at will.

3.1.2.3 Subjective Data: Prioritizing

SMEs rated how well participants set and used priorities. A significant main effect of condition was found on marking flight strips while performing other tasks and overall prioritizing, $F(3, 21) = 3.99$, $p = .021$, and $F(3, 18) = 4.40$, $p = .017$, respectively. SMEs rated the participants as marking flight strips less well in the D/NI condition than in the B and D conditions. SMEs also rated overall prioritizing in the D/NI condition lower than in the B condition, as shown in Figure 2. The remaining three items in this section of the SME rating form (taking actions in an appropriate order, preplanning control actions, and handling control tasks for several aircraft) did not show any significant difference between conditions.
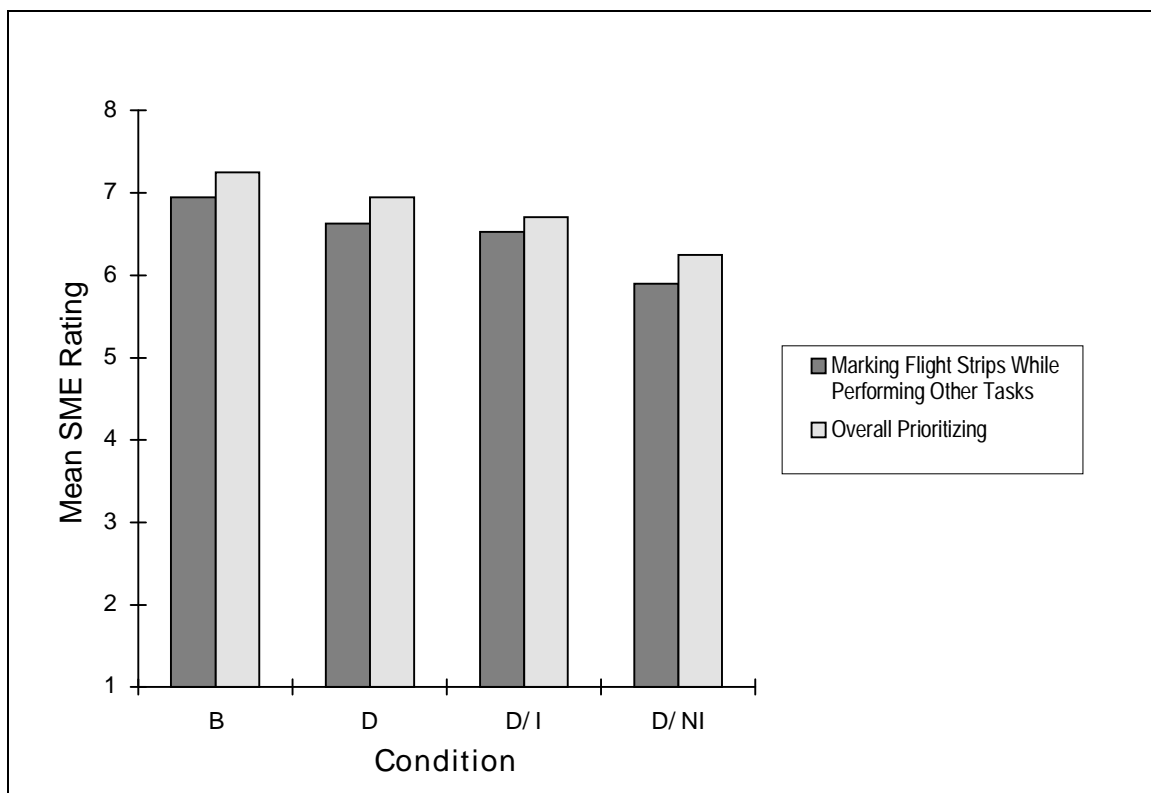

Figure 2. Prioritization of tasks.

3.1.3 Control Strategies

The way controllers altered their strategies for controlling traffic under free flight was also examined.

3.1.3.1 Altitude Changes

The number of altitude changes made in each experimental run was recorded by ATCoach. No significant difference was found between the four conditions in the number of altitude changes

made. This result is consistent with ratings given by the SMEs on the Observation Form who did not indicate that the participants' preference for vertical separation was significantly different among the four conditions.

3.1.3.2 Speed Changes

A significant main effect of condition was found on the number of speed changes made by the participants, $F(3, 24) = 5.46$, $p = .005$. Participants made significantly more speed changes in the B condition than in the three free flight conditions (see Figure 3). This result is consistent with SME ratings on the preference for speed control item on the Observation Form where a significant main effect of condition was found, $F(3, 21) = 3.65$, $p = .004$. SMEs rated the participants as preferring speed control more in the B condition than in D/I and D/NI conditions (see Figure 4).

These results suggest that speed changes may be less useful in a free flight environment. This can be understood given how controllers use speed changes. In the current system, controllers mainly use speed changes to separate aircraft that are proceeding in trail along an airway. If the leading aircraft is slower than the trailing aircraft, a loss of separation may occur as the trailing aircraft overtakes the leading aircraft. To prevent this, controllers speed up the leading aircraft or slow down the trailing aircraft so that separation remains constant. In a free flight system where airways are not followed, situations like this will probably occur less frequently, resulting in fewer requirements for speed changes.

3.1.3.3 Heading Changes

The number of heading changes made in each scenario was recorded by ATCoach. No significant difference was found between the four conditions in the number of heading changes made. This result is consistent with ratings given by the SMEs on the Observation Form who did not rate the participants' preference for vectoring differently between the four conditions.

These results indicate that participants instituted almost the same number of altitude and heading changes as a result of the free flight conditions examined, although they did show a reduction in the use of speed as a control strategy.

3.1.4 Communications

Data regarding the communications between the participants and the pseudopilots was derived from two sources, the push-to-talk (PTT) actions recorded by the communication system and SME ratings made on the Observation Form.
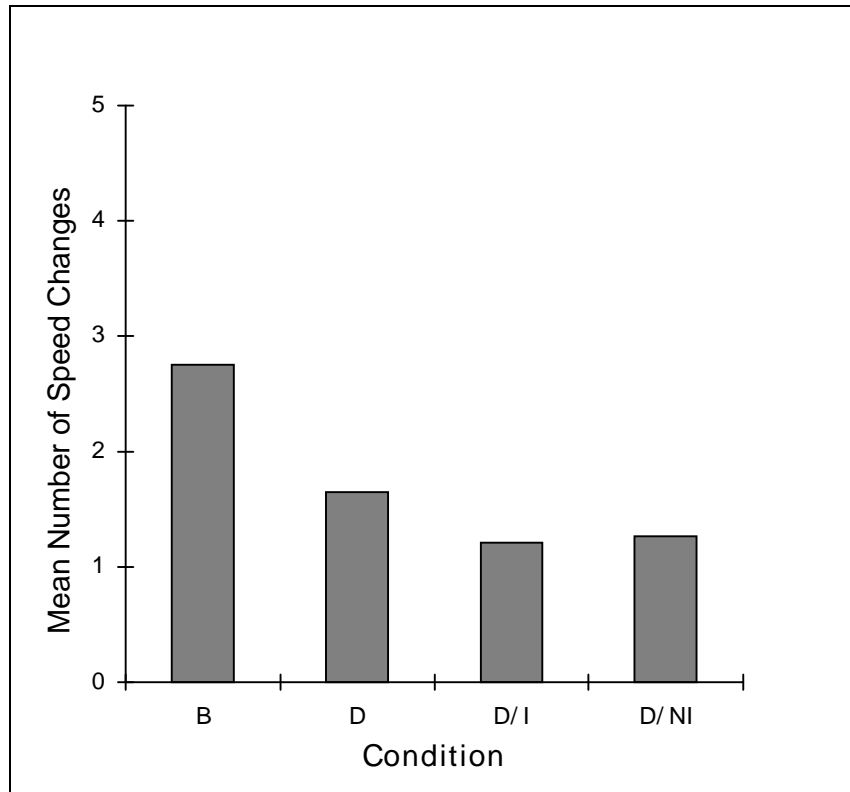
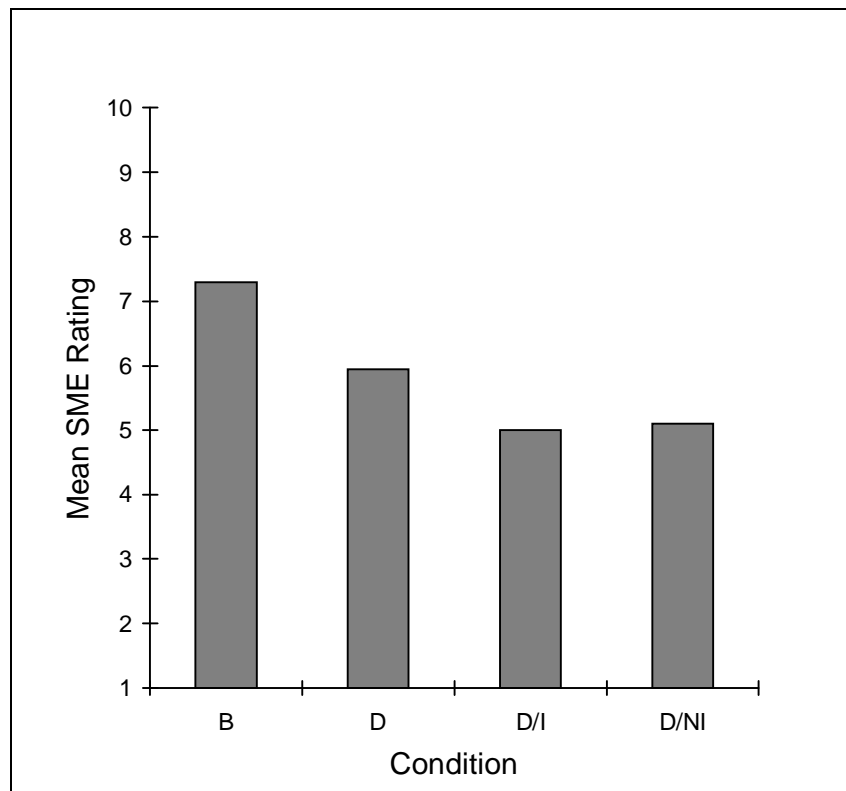Figure 3. Number of speed changes issued.

Figure 4. Controller's preference for speed control.

3.1.4.1 Objective Data: Push-to-Talk Communications

The simulation system counted the number and duration of communications made by the participants to the pseudopilots. No significant differences were found between the four conditions in the number of PTT communications or in the duration of the communications.

3.1.4.2 Subjective Data: Communication

SMEs rated how well participants communicated with the pseudopilots using the Communicating section of the Observation Form. No significant difference was found between four conditions on the four items included in this section: using proper phraseology, communicating clearly and efficiently, listening for pilot readbacks and requests, and overall communicating.

3.1.4.3 Subjective Data: Providing Control Information

SMEs rated how often participants provided essential and additional ATC information to the pilots using the Providing Control Information section of the Observation Form. A significant main effect of condition was found on the providing essential ATC information item, $F(3, 21) = 4.11$, $p = .019$. SMEs rated the participants as providing significantly less essential ATC information in the D/NI condition than in the B condition. A significant main effect of condition was also found on the providing additional ATC information item, $F(3, 21) = 3.65$, $p = .029$. SMEs rated the participants as providing significantly less additional ATC information in the D/NI condition than in the B condition (see Figure 5). However, the overall providing control information item did not show a significant difference between conditions. While this analysis does not show a greater frequency of communications occurring or changes in communication quality, it does indicate changes in communication content. Controllers apparently provided less information (both essential and additional information) under the higher levels of free flight.

3.1.5 Overall Controller Performance

Both the participants and SMEs rated how well the participants controlled traffic during the scenario. Participants made these ratings (regarding their own performance) on the Post-Scenario Questionnaire and SMEs made these ratings on the Observation Form. A significant main effect of condition was not found in the participants' ratings of their own performance, $F(3, 12) = 0.11$, $p = .955$, but was found in the SME ratings of participant performance, $F(3, 21) = 5.50$, $p = .001$. (Differences in degrees of freedom on this test reflect missing data points on some participants' questionnaires.) SMEs rated the participants' performance in the D/I and D/NI conditions as significantly lower than in the B condition (see Figure 6). No reliable correlation was found between the ratings given by the participants and SMEs on this item, $r(67) = .08$, $p = .513$. Participant and SME performance ratings also did not significantly correlate with the number of conflicts that occurred during the run. Therefore, it would appear that these two measures tapped into independent factors concerning performance and traffic separation.
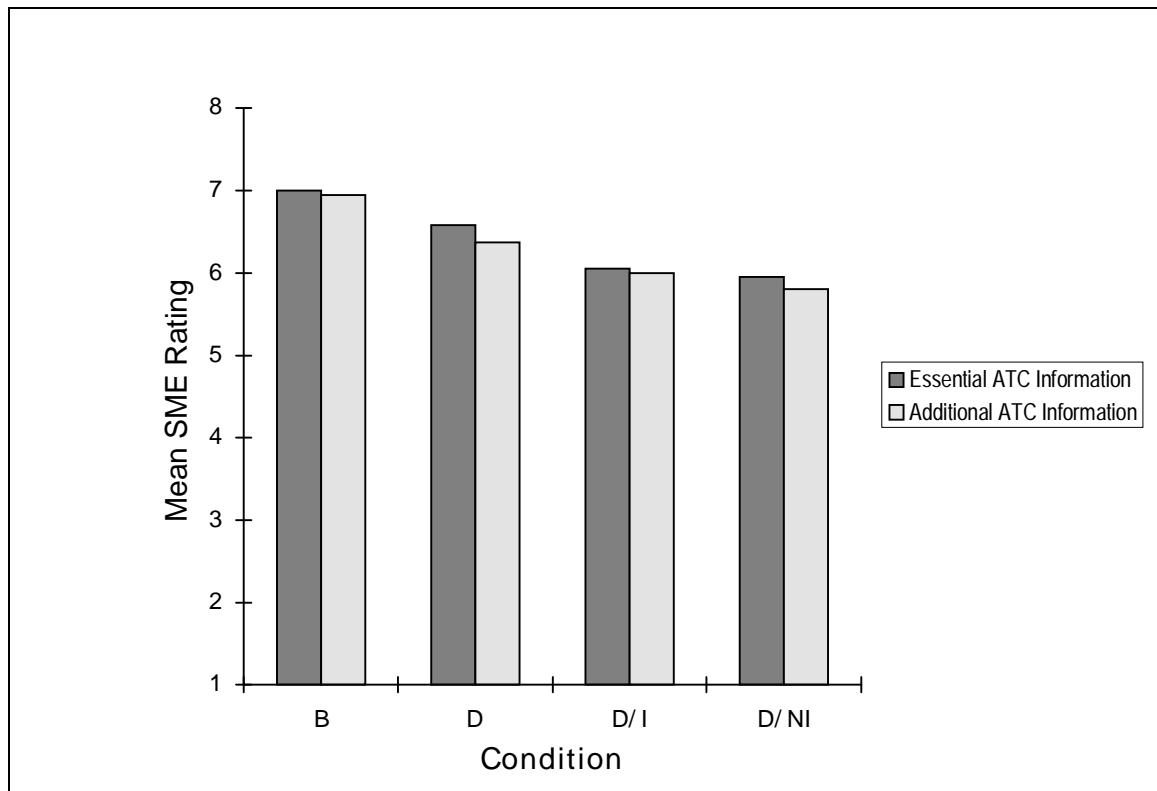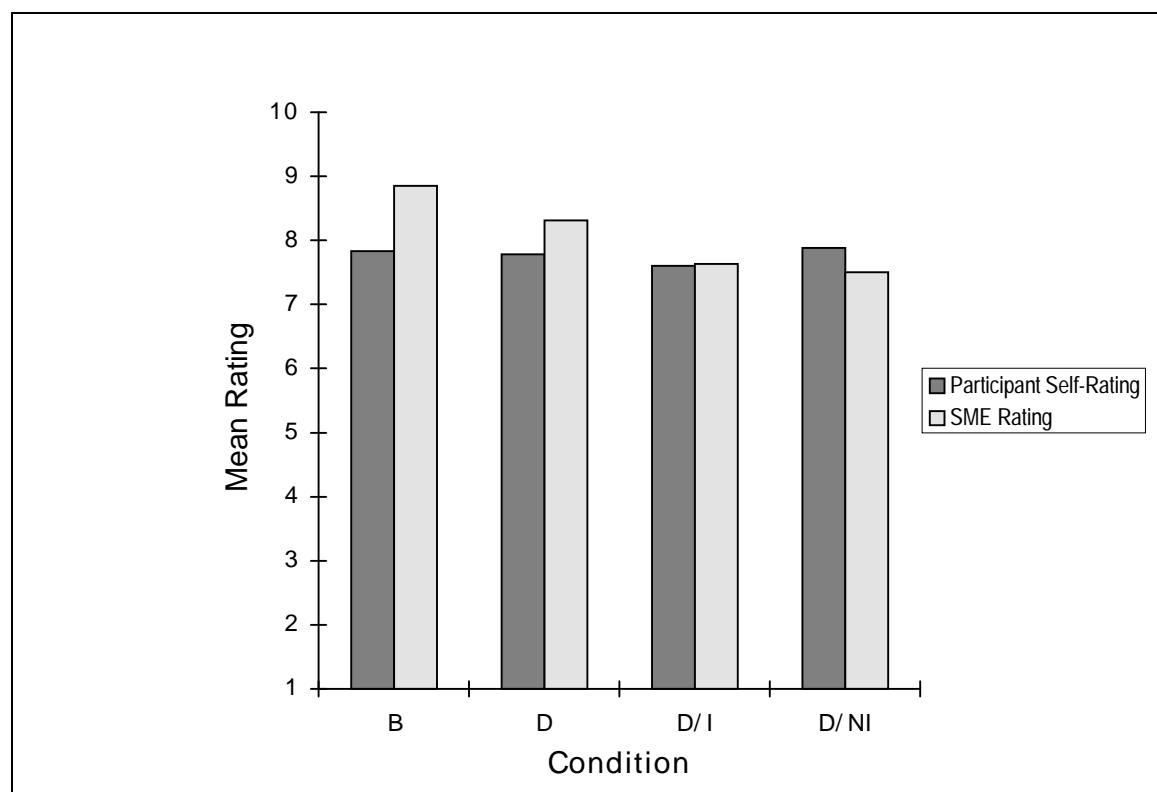
Figure 5. Providing control information.

Figure 6. Ratings of overall performance.

3.2 Controller Workload

Both physical task load and estimates of subjective mental workload were obtained.

3.2.1 Objective Data: Task Load

The number of data and slew entries made by the participants was recorded by ATCoach. No significant difference was found between the four conditions in the number of data and slew entries made.

3.2.2 Subjective Data: Workload

3.2.2.1 End-of-Trial SME Ratings and Controller Self Ratings

At the conclusion of each experimental run, participants and SMEs rated how hard the participant had worked during the scenario. Participants made these ratings on the Post-Scenario Questionnaire and SMEs made these ratings on the Observation Form. A significant main effect of condition was found for self-ratings and SME ratings of workload, $F(3, 21) = 6.08$, $p = .001$ and $F(3, 21) = 7.79$, $p = .006$ respectively. SMEs rated workload as significantly lower in the B condition than in the three free flight conditions. Controllers rated workload as significantly lower in the B condition than the D/I and D/NI conditions (see Figure 7). There was a significant positive correlation (although low) between the ratings given by the participants and SMEs on this item, $r(77) = .32$, $p = .005$.
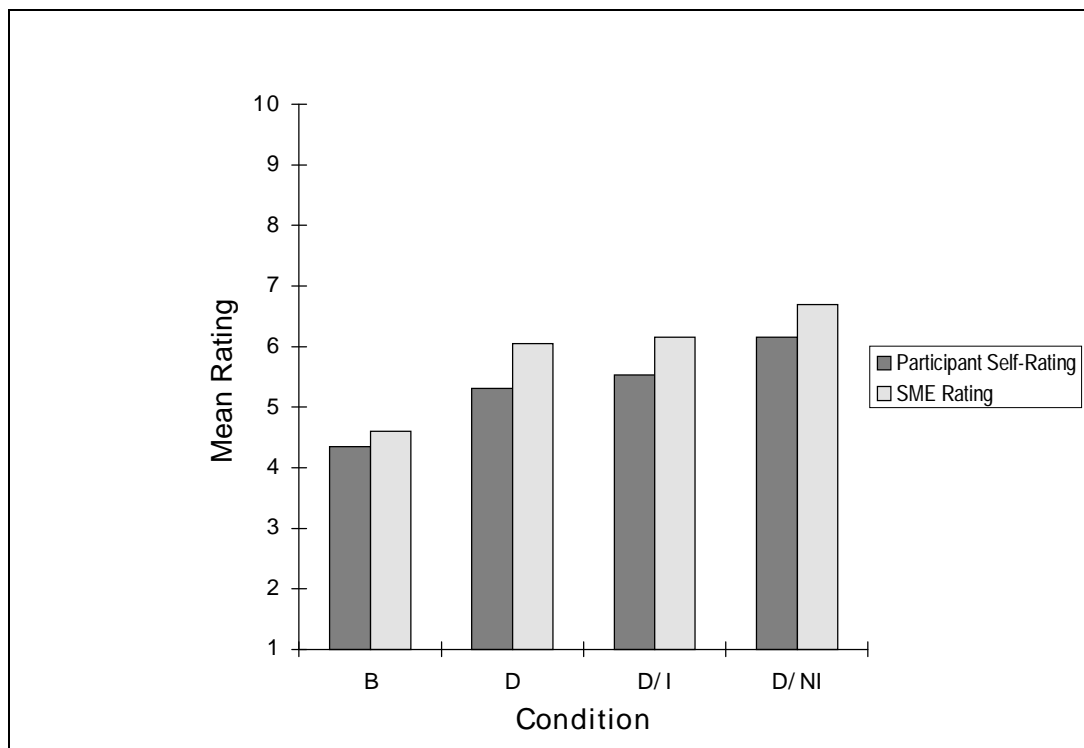

Figure 7. Subjective workload ratings.

19

3.2.2.2 ATWIT Ratings

ATWIT ratings did not show a statistically significant difference between conditions, $F(3, 27) = 1.64$, $p = .204$. The B condition received ATWIT ratings approximately 7% lower than the three free flight conditions, as shown in Figure 8.
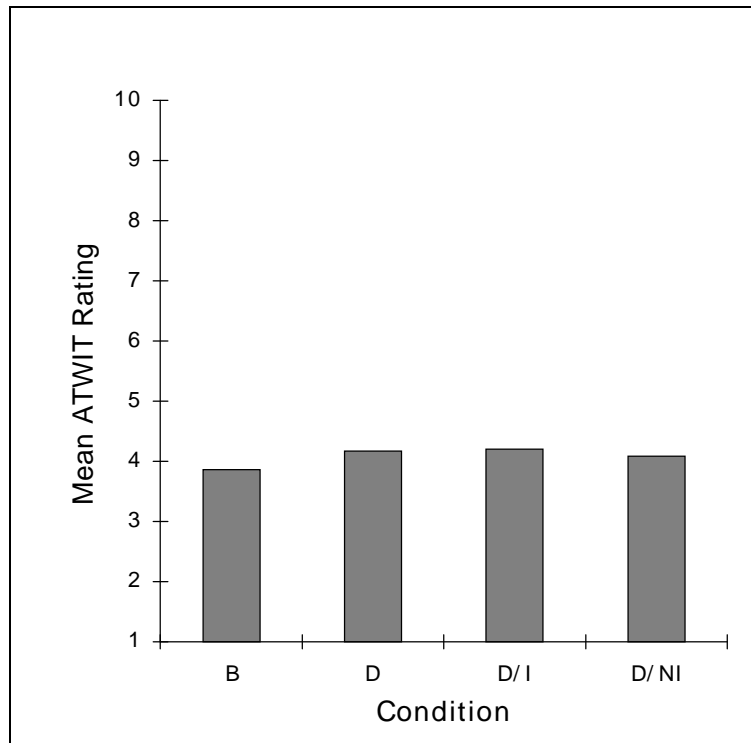
Figure 8. ATWIT ratings of workload.

3.2.2.3 NASA-Task Load Index Ratings

The participants' ratings on the NASA-TLX scales of mental load, physical load, effort, temporal load, stress, and frustration were weighted based on each participants' rankings of these sub-scales. Their resultant overall NASA-TLX scores were significantly different between conditions $F(3, 66) = 9.07$, $p = .000$. The scores for the B condition were significantly lower than for the three free flight conditions as illustrated in Figure 9.

3.2.2.4 Inter-Correlation of Workload Measures

Because the ATWIT ratings did not show a significant effect of condition, additional analyses were performed to examine how the ATWIT ratings related to the post-scenario workload ratings. The mean ATWIT rating for the trial showed moderate correlations with participants' self-rating of workload, $r(78) = .54$, $p = .000$, and the NASA-TLX scores, $r(76) = .52$, $p = .000$. A smaller correlation with the SME's rating of workload was found, $r(78) = .27$, $p = .015$. It was hypothesized that the post-scenario ratings may have been more reflective of workload at the end
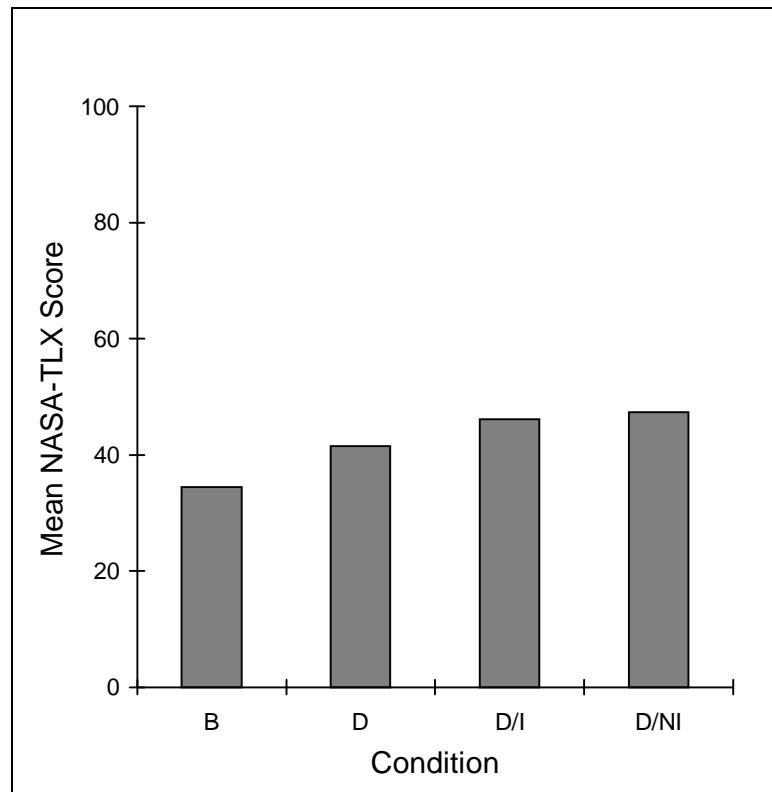
Figure 9. NASA-TLX scores.

of the run or during workload peaks as opposed to workload over the entire run. However, this was not the case. The average ATWIT rating across the trial correlated with the post-scenario workload ratings as well or better than any subset of the ATWIT ratings including the highest rating, the lowest rating, the last rating, and the mean of the last three ratings for the scenario.

3.3 Situation Awareness

The SMEs provided several subjective ratings related to SA and the SAGAT data provided a measure of participant SA.

3.3.1 Subjective Data: SME Ratings of Maintaining Attention and Situation Awareness

SMEs rated participant SA using the Maintaining Attention and Situation Awareness section of the Observation Form. A significant main effect of condition was found on two items: detecting pilot deviations from control instruction, $F(3, 15) = 7.83$, $p = .002$, and overall attention and situation awareness, $F(3, 21) = 3.82$, $p = .025$. SMEs rated the participants as detecting pilot deviations better in the B condition than in the D/NI condition. (Although what they considered a deviation under the D/NI condition was difficult to ascertain.) SMEs rated the participants as showing significantly higher overall attention and SA in the B condition than in the D/NI condition (see Figure 10). The remaining three items (maintaining awareness of aircraft positions, ensuring positive control, and correcting own errors in a timely manner) did not show any significant difference between conditions.

21
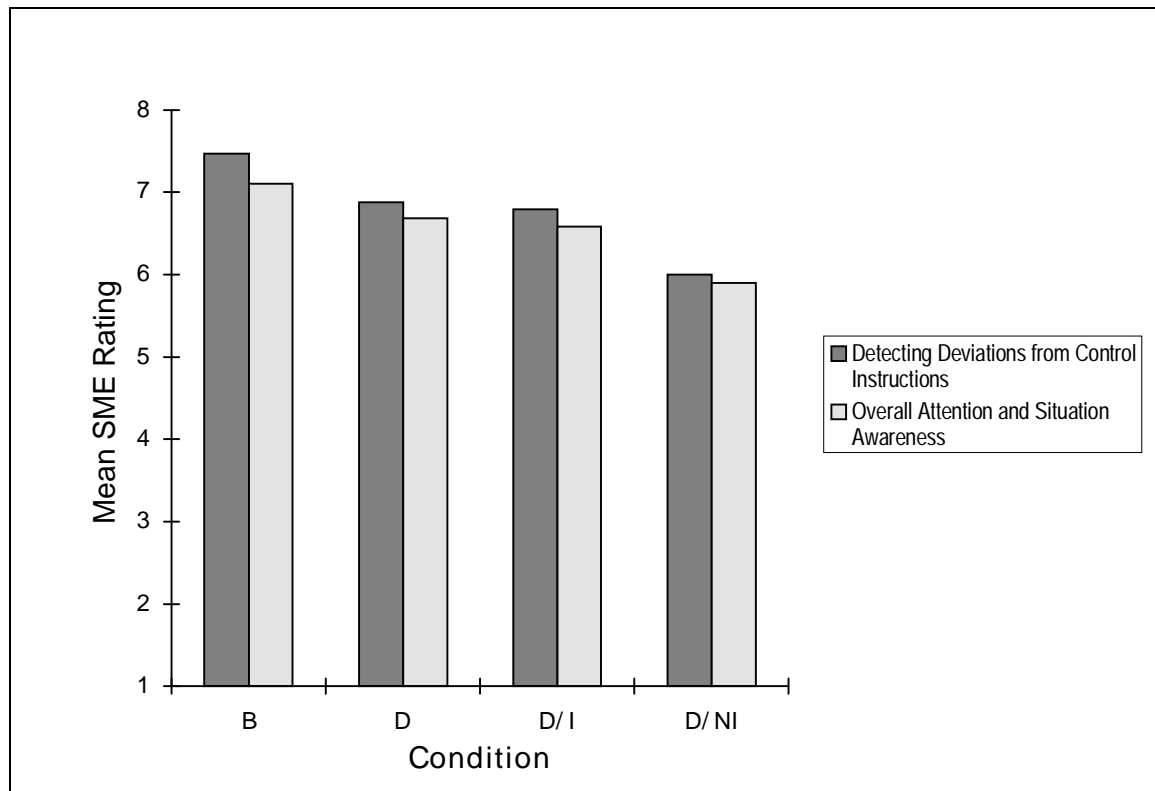
Figure 10. SME ratings of attention and situation awareness

3.3.2 SAGAT Results

The participants' perception of the traffic situation as reported on the SAGAT queries were compared to the actual state of the traffic situation at the time of each freeze. Their answers were scored as correct or incorrect and subjected to an arcsine transformation (to correct for non-normality of binomial data).

3.3.2.1 Level 1 SA- Perception of the Traffic Elements

Of the Level 1 SA queries, participants' knowledge of the presence and location of aircraft at the time of the freeze (+/- 5 mi) was significantly different between conditions, $F(3, 298) = 3.20$, $p = .024$. Participants were aware of significantly fewer aircraft in the D/I and D/NI conditions than in the B condition, shown in Figure 11. This was true for all aircraft in the sector (including those in active control, those recently handed-off, and those soon to be handed-off).

A significant difference was found when examining participants' awareness of only the aircraft currently in their active control, $F(3, 298) = 2.72$, $p = .045$. Participants were aware of significantly fewer active aircraft in the D condition than in the B condition.

Figure 11. Awareness of aircraft location.
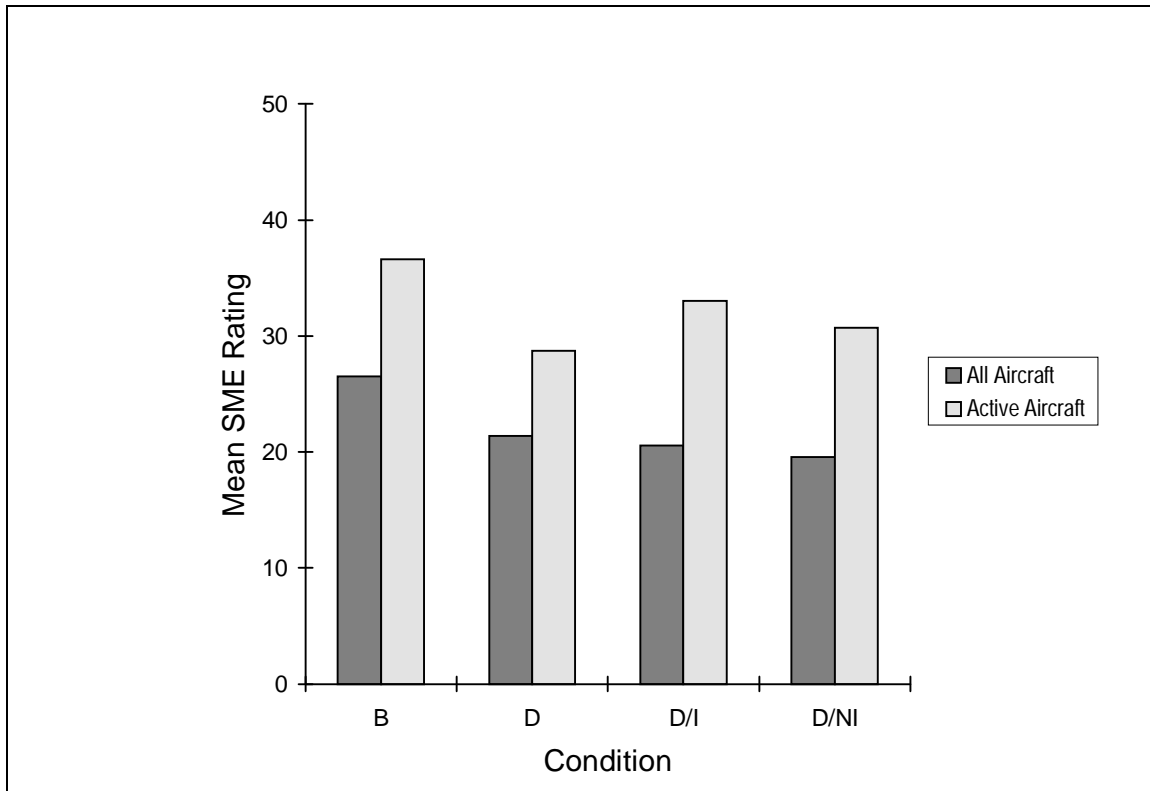
For those aircraft of which participants were aware, the remaining SAGAT queries were posed. Of the Level 1 SA queries, only the participants' awareness of the alphabetic portion of aircraft callsigns (the alphabetic and numeric portions of the callsign were scored separately), was significantly different between conditions, $F(3, 293) = 3.70$, $p = .012$. As shown in Figure 12, participants were able to report on aircraft callsigns significantly less in the D and D/NI conditions. None of the other Level 1 SA queries were significantly different between conditions including: level of control, callsign (numeric portion), altitude, groundspeed, heading, vertical change, turning, or aircraft type.

3.3.2.2 Level 2 and 3 Situation Awareness - Comprehension and Projection of Traffic Situation

A number of the queries pertaining to the participants' comprehension of what happened in the traffic scenarios and their ability to project the actions of the traffic were significantly impacted by the free flight conditions tested. Participants' ability to identify the next sector to which an aircraft would transition (indicating an awareness of future flight path/direction) was significantly impacted by the free flight conditions tested, $F(3, 264) = 3.18$, $p = .025$. As shown in Figure 13, participants were significantly more aware of this information in the B condition than in the three free flight conditions. Participants' ability to report on which aircraft were being impacted by weather (or would be in the next 5 minutes) was also different across the four conditions, $F(3, 178) = 4.76$, $p = .003$. Awareness of weather impact was significantly lower in the D and the D/NI conditions as compared to the B condition, as shown in Figure 14.
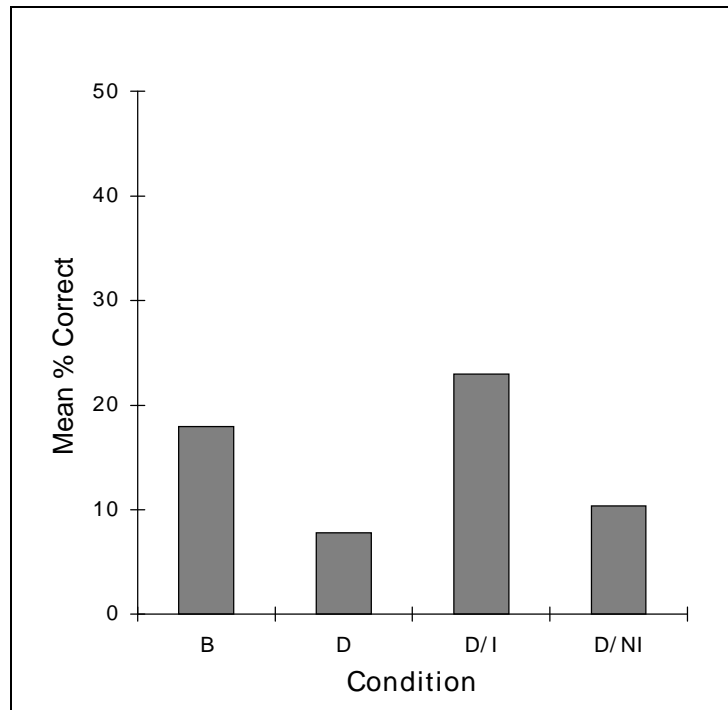
23

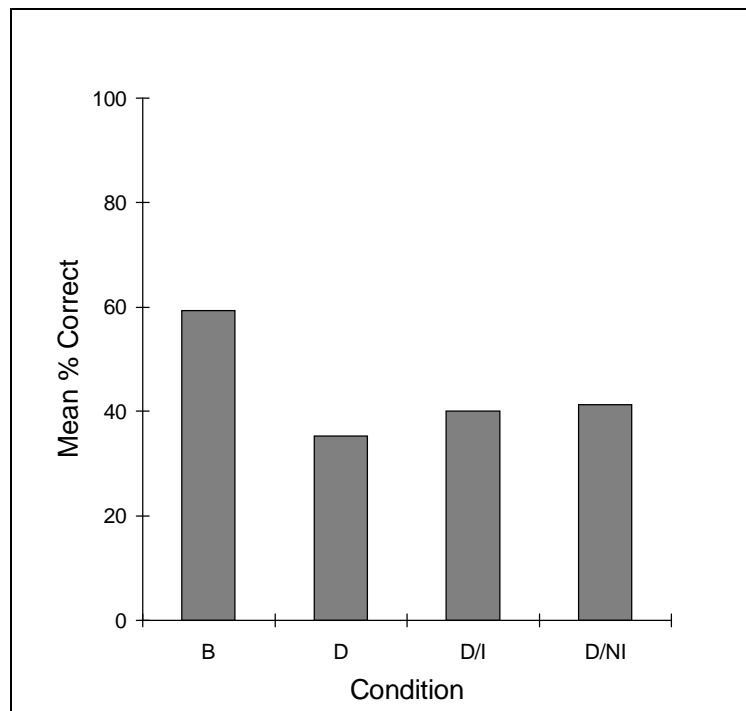Figure 12. Awareness of aircraft callsign (alphabetic portion).

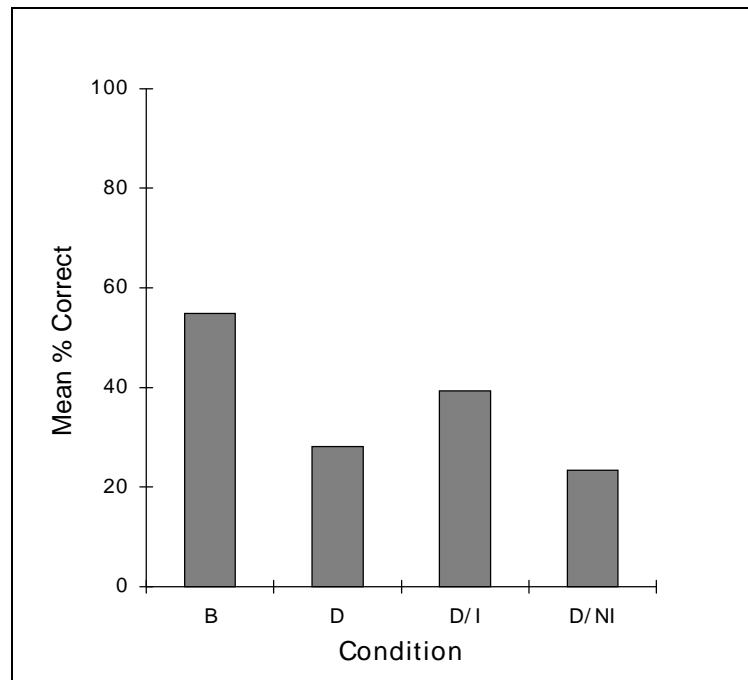Figure 13. Awareness of aircraft next sector.

Figure 14. Awareness of weather impact.

Although not significant at the .05 level, several other SA variables were different across conditions at levels that approached statistical significance. These variables are as follows: the participants' ability to identify aircraft with incomplete clearances, $F(3, 280) = 2.29$, $p = .079$, aircraft that had correctly received the clearance, $F(3, 193) = 2.55$, $p = .057$, and aircraft that were conforming to their clearances, $F(3, 193) = 2.55$, $p = .057$.

For these questions, the trends indicate that participants were somewhat less aware of aircraft which were in a transition state (responding to clearance changes), as shown in Figures 15, 16, and 17. This result is in agreement with the SMEs' lower subjective rating for detecting deviations from control instructions under free flight conditions. Overall, it would appear that controller SA was lower under the free flight conditions in terms of their ability to keep up with the traffic and to predict its actions over time.

3.4 Simulation Fidelity Evaluation

As a check on the quality and veracity of the simulation used to obtain these results, several scales on the participant's Post-Scenario Questionnaire and the Exit Questionnaire addressed the realism and fidelity of the simulation. On the Post-Scenario Questionnaire, a significant main effect of condition was found, $F(3, 21) = 7.74$, $p = .001$. Controller participants rated scenarios in the D/NI condition as significantly less realistic than scenarios in the B and D conditions (see Figure 18). This may reflect the novelty of such a condition for the controllers. No significant difference was found in realism rating between the 14 scenarios independent of condition, $F(13, 64) = 1.15$, $p = .339$. On the Exit Questionnaire, participants rated the simulations with a mean rating of 6.4 on a 1 to 10 scale (with 1 being extremely unrealistic and 10 being extremely realistic).
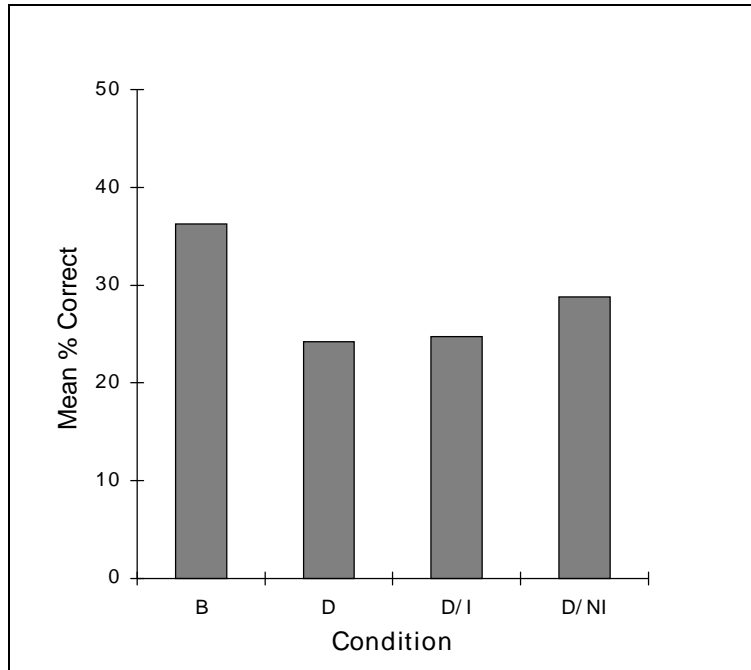
25

Figure 15. Awareness of aircraft with incomplete clearance changes.
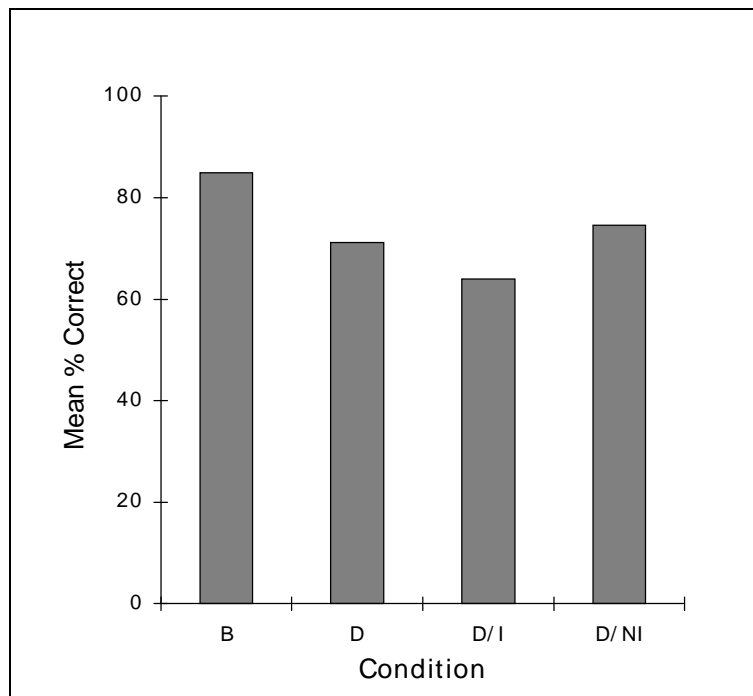

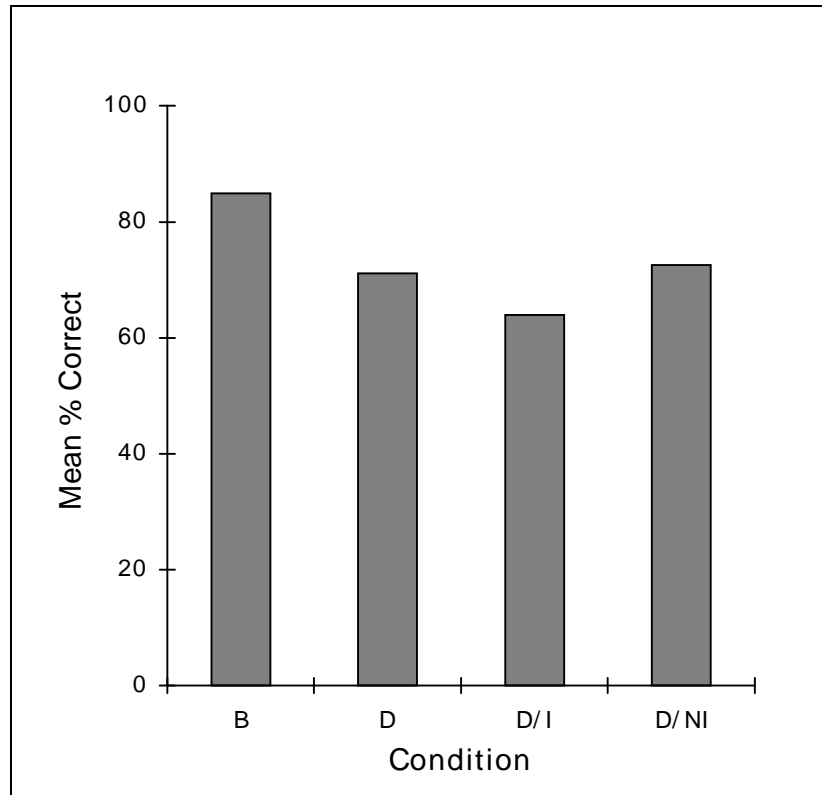Figure 16. Awareness if clearance was received correctly by correct aircraft.

Figure 17. Awareness if aircraft is conforming to clearance.
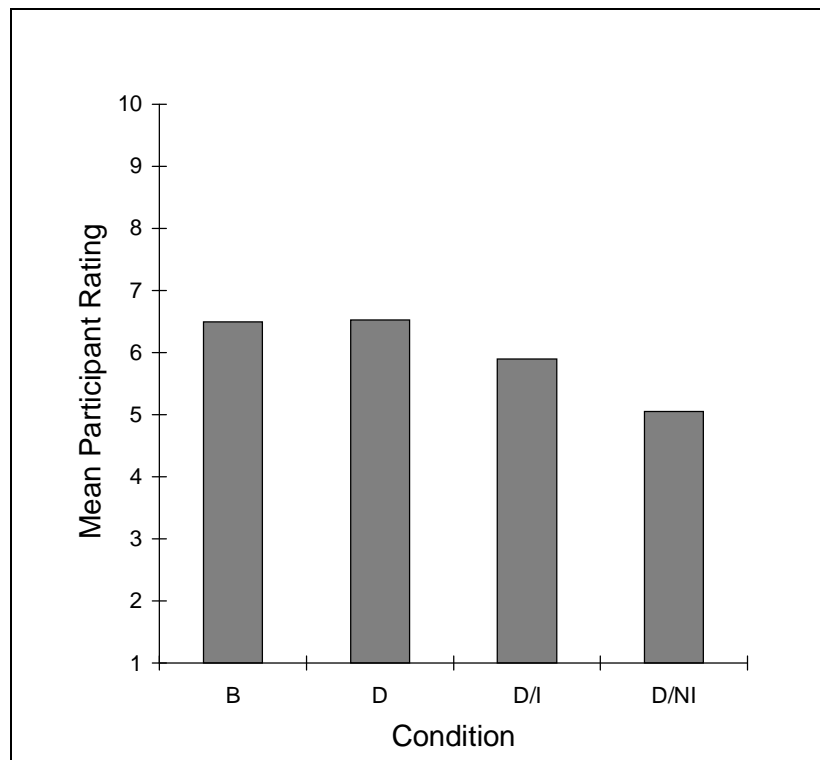

Figure 18. Simulation realism ratings.

Participants also rated the degree to which the ATWIT device interfered with their normal ATC operations. No significant difference was found between the conditions on this item on the Post-Scenario Questionnaire. On the Exit Questionnaire, participants rated the ATWIT as interfering with a mean of 3.1 on a 1 to 10 scale (with 1 being no interference and 10 being extreme interference), indicating a fairly low level of interference.

Participants rated how well the pseudopilots responded to clearances and callbacks. No significant difference was found between the conditions on this item on the Post-Scenario Questionnaire. On the Exit Questionnaire, participants rated the pseudopilots with a mean score of 8.3 on a 1 to 10 scale (with 1 being not adequate and 10 being adequate), indicating a fairly high assessment of pseudopilot performance during this study.

Participants rated the adequacy of the training runs on the Exit Questionnaire with a mean rating of 7.0 on a 1 to 10 scale (with 1 being not adequate and 10 being adequate), indicating they were reasonably adequate.

4. Discussion

This study investigated the effects of free flight on controllers without new supporting technologies. It did not investigate the SA of pilots who would be operating under free flight. It may, however, provide an indication of the difficulties that controllers could have in detecting and preventing a loss of aircraft separation under some possible free flight conditions. This study indicates that in those cases where the pilots have failed to separate themselves, controllers may have problems detecting and preventing loss of separation as their role changes from that of an active controller to one of traffic monitor. This statement is predicated on several observations.

While sensitive measures of controller performance are difficult to find under the constraints of simulation testing, a trend towards greater loss of separation was observed in the highest level of free flight, which allowed pilots to deviate at will. In addition, several performance measures indicated that there was a tendency to fall behind in prioritizing and performing tasks. This observation was accompanied by an increase in mental workload with succeeding levels of free flight. The loss of organization provided by the normal route structure due to the use of direct routing is a possible factor driving this workload increase. In our simulation, it is likely that the controllers had to work harder to keep up with what aircraft were doing instead of relying on the typical patterns that the airways provide. In those conditions where the pilots could also deviate at will, controller workload increased even more. The lower predictability of the free flight conditions may have resulted in a need for controllers to expend more effort in monitoring the traffic situation. They did not exhibit a pattern of less workload, which might be assumed to be the case for a monitoring situation.

Controller SA was also negatively impacted by the free flight conditions tested. SA is critical for detecting separation problems and managing air traffic. Without it, controllers will be ineffective as monitors of even very sophisticated automated systems. The ability of controllers to maintain an up-to-date picture of a dynamic and complex traffic situation depends on their ability to integrate a great deal of data on many aircraft into an internal structure that allows relationships between aircraft to be understood (e.g., which aircraft are traffic for each other). As the

predictability of the aircraft decreases, this becomes a much more difficult job. More of the controllers' limited attention and working memory is required to process each aircraft. As a result, the participants in this study were unable to maintain awareness of as many aircraft under the free flight conditions. Although this may be acceptable as long as they were the important aircraft, it has been shown that errors may occur when the controller's attention is directed towards other traffic situations that are believed to be important but may not be (Endsley & Rodgers, 1996; Jones & Endsley, 1996). A reduced awareness of aircraft in the traffic situation increases the probability that a loss of separation will occur.

In addition, the higher levels of SA (comprehension and projection of the situation) were negatively impacted. In particular, participants showed a decreased understanding of the projected paths of the aircraft (reflecting the lower predictability of the free flight conditions). They also showed a reduced ability to assess what was happening with the aircraft (in terms of insuring conformance to clearances and weather impacts that might induce the need for deviations from flight plans). Interestingly, decreases in the higher levels of SA have also been observed in several studies of automation (Carmody & Gluckman, 1993; Endsley & Kaber, in review; Endsley & Kiris, 1995a). When people become passive monitors of information rather than active processors (watching a traffic scenario rather than creating their own plan for controlling the traffic flow), SA can decrease. This factor has been directly linked to the out-of-the-loop performance problem in which monitors of automated systems are slow or unable to detect and intervene during failures of automated systems (Endsley & Kiris, 1995b). In this study, it is likely that a similar phenomenon occurred, leading the controllers to have a reduced understanding of the traffic picture as active control decreased under free flight conditions. Similarly, by virtue of being out-of-the-loop, they were less likely to be able to intervene to insure separation. Under sustained vigilance conditions, it is probable that this problem will be even greater.

It may be that the higher level of mental workload led to lower SA. Conversely, one could postulate that the controllers were concerned about their SA and, therefore, expended more effort (higher workload) to correct it. A more parsimonious explanation, however, is that the same factor drove both the increased workload and reduced SA: the lower predictability of the air traffic situation created by the use of direct routing and the ability of the pilots to deviate from filed flight plans at will.

While a higher number of operational errors was only observed in the condition in which pilots could deviate at will without informing the controllers of their intentions, higher workload and lower SA were observed across all the free flight conditions. In an operational setting, particularly if vigilance problems occur as a result of monitoring conditions, a decrease in controller performance is likely to result from lower SA and higher workload.

This finding does not, however, mean that free flight is infeasible. Rather, it indicates a critical factor that will need to be overcome or compensated for if controllers are to remain as an effective part of the future ATC system. If displays, automated systems, or procedures can be developed for assisting controllers in regaining some degree of predictability, it is expected that much of this problem can be resolved. Increased predictability of air traffic should significantly

assist controllers in developing the mental picture needed at a lower level of workload. As an example, increased predictability might be achieved by drawing on data link to create informed displays. Simple display management tools, which allow the controller to manage the complexity (lack of predictability) of the traffic flows (through highlighting, coding, or projection activities), might be effective. The use of graphic displays for portraying the higher levels of SA should specifically be explored to compensate for losses that may occur under passive processing. Whichever concepts are explored, the development of these systems will need to proceed very carefully to insure that they provide needed information without creating display overload.

While conflict alert technologies might be effective as a backup aid, display technologies that allow the controller to function well independently should also be explored. Historically, people's ability to respond to alarms has been problematic. High false alarm rates produce mistrust, which can lead to ignoring real problems. High hit rates can produce complacency and lack of vigilance. Both of these conditions can lead to serious errors. By keeping the controllers alert, informed, and involved in the situation, they will be able to function effectively as a much needed safety net even in the highest levels of free flight. This can be accomplished by providing them with the information they need to perform effectively.

The participants in this study were only exposed to each free flight condition for a few hours. Therefore, it may be the case that, with further practice, some of the problems identified in this report will diminish. Controllers may develop compensatory strategies to cope with the changes in procedures and responsibilities inherent in free flight. For example, in the D/N and D/NI conditions, some study participants learned to query aircraft entering the sector about their intentions. In this way, they were able to better anticipate independent pilot actions responding to weather and turbulence. While this approach may not be ideal, it illustrates an initial attempt by controllers to adjust to the new situation. As controllers spend more time in a free flight environment, they may adapt their information requirements to their new responsibilities. It is still the case that they will need to maintain some level of SA to be able to intervene if conflicts develop. However, studies that permit time to acclimate to a free flight environment will be needed before final conclusions can be drawn regarding the requirement for displays or procedures for augmenting aircraft predictability.

5. Conclusions

In examining the aspects of free flight, this study showed that if controllers are expected to act as passive monitors of free flight air traffic, their awareness of the state of air traffic may be reduced, their workload may increase, and their ability to intervene in a timely manner may be limited. If the future air traffic system is able to function autonomously (i.e., through automation or pilot-insured separation), this may not be a problem. If, however, the controller will be retained in the system as safety net for insuring that separation between aircraft is not lost, for negotiating disputes, or for easing difficult or congested transitions, it is critical that solutions are generated for these problems. Intervening technologies need to be explored that will redress the loss of SA that can occur even with low levels of free flight. In addition, alternate operational

concepts might be examined that provide some of the benefits of free flight while allowing the controller to maintain an active role in aircraft separation.

This study is one of the first conducted to look systematically at free flight to discover the actual implications of such concepts on the ability of its participants to function in a safe and efficient manner. As such, it was fairly limited in the scope of free flight concepts examined. It did, however, reveal that some of the concerns about the effects of free flight are well founded and should be taken into account to inform the development of systems for free flight. Specifically, it highlighted the critical role that predictability plays in allowing controllers to develop an accurate and complete picture of the air traffic situation. The development of compensating technologies and strategies needs to be explored for dealing with this issue. In addition, more studies are needed that will expand this effort to collect objective data on other aspects of free flight and its consequences on the air transportation system. Only in this way is it likely that the needed enabling technologies and mechanisms will be developed for allowing a safe free flight to become reality.

References

ATCoach (Version 7.0) [Computer software]. (1996). Lexington, MA: UFA, Inc.

Carmody, M. A., & Gluckman, J. P. (1993). Task specific effects of automation and automation failure on performance, workload and situational awareness. In R. S. Jensen & D. Neumeister (Ed.), *Proceedings of the Seventh International Symposium on Aviation Psychology* (pp. 167-171). Columbus: Department of Aviation, The Ohio State University.

Endsley, M. R., & Kaber, D. B (1997). *Level of automation effects on performance, situation awareness and workload in a dynamic control task*. Manuscript submitted for publication.

Endsley, M. R., & Kiris, E. O. (1995a). *Situation awareness global assessment technique (SAGAT) TRACON air traffic control version user guide*. Lubbock: Texas Technical University.

Endsley, M. R., & Kiris, E. O. (1995b). The out-of-the-loop performance problem and level of control in automation. *Human Factors, 37*(2), 381-394.

Endsley, M. R., & Rodgers, M. D. (1996). Attention distribution and situation awareness in air traffic control. In *Proceedings of the 40th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 82-85). Santa Monica, CA: Human Factors and Ergonomics Society.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139-183). Amsterdam: North-Holland.

Jones, D. G., & Endsley, M. R. (1996). Sources of situation awareness errors in aviation. *Aviation, Space and Environmental Medicine, 67*(6), 507-512.

Sollenberger, R. L., Stein, E. S., & Gromelski, S. (1997). *The development and evaluation of a behaviorally based rating form for assessing air traffic controller performance* (DOT/FAA/CT-TN96/16). Atlantic City, NJ: Federal Aviation Administration Technical Center.

Stein, E. S. (1985). *Air traffic controller workload: An examination of a workload probe* (DOT/FAA/CT-TN84/24). Atlantic City, NJ: Federal Aviation Administration Technical Center.

RTCA. (1995). *Report of the RTCA Board of Directors select committee on free flight.* Washington, DC: Author.

Stein, E., and Guttman, J. A. (1997). *The influence of generic airspace on en route air traffic controller performance.* Manuscript in preparation, Atlantic City, NJ: DOT/FAA William J. Hughes Technical Center.

Instructions for questions 1-24

This form was designed to be used by instructor-certified air traffic control specialists to evaluate the effectiveness of controllers working in simulation environments. Observers will rate the effectiveness of controllers in several different performance areas using the scale shown below. When making your ratings, please try to use the entire scale range as much as possible. You are encouraged to write down observations, and you may make preliminary ratings during the course of the scenario. However, we recommend that you wait until the scenario is finished before making your final ratings. The observations you make do not need to be restricted to the performance areas covered in this form and may include other areas that you think are important. Also, please write down any comments that may improve this evaluation form. Your identity will remain anonymous, so do not write your name on the form.

Rating	Label Description
1	Controller demonstrated extremely poor judgment in making control decisions and very frequently made errors.
2	Controller demonstrated poor judgment in making some control decisions and occasionally made errors.
3	Controller make questionable decisions using poor control techniques which led to restricting the normal traffic flow.
4	Controller demonstrated the ability to keep aircraft separated but used spacing and separation criteria which was excessive.
5	Controller demonstrated adequate judgment in making control decisions.
6	Controller demonstrated good judgment in making control decisions using efficient control techniques.
7	Controller frequently demonstrated excellent judgment in making control decisions using extremely good control techniques.
8	Controller always demonstrated excellent judgment in making even the most difficult control decisions while using outstanding control techniques.

Maintaining Safe and Efficient Traffic Flow

	1	2	3	4	5	6	7	8

1. Maintaining Separation and Resolving Potential Conflicts
   - using control instructions that maintain safe aircraft separation
   - detecting and resolving impending conflicts early
 Comments:								

2. Sequencing Arrival and Departure Aircraft Efficiently
   - using efficient and orderly spacing techniques for arrival and departure aircraft
   - maintaining safe arrival and departure intervals that minimize delays
 Comments:								

3. Using Control Instructions Effectively
   - providing accurate navigational assistance to pilots
   - avoiding clearances that result in the need for additional instructions to handle aircraft completely
   - avoiding excessive vectoring or over-controlling
 Comments:								

4. Overall Safe and Efficient Traffic Flow Scale Rating **Comments:**								

Maintaining Attention and Situation Awareness

5. Maintaining Awareness of Aircraft Positions
  - avoiding fixation on one area of the radar scope when other areas need attention
  - using scanning patterns that monitor all aircraft on the radar scope
Comments:
	1	2	3	4	5	6	7	8

6. Ensuring Positive Control **Comments:**	1	2	3	4	5	6	7	8

7. Detecting Pilot Deviations from Control Instructions
  - ensuring that pilots follow assigned clearances correctly
  - correcting pilot deviations in a timely manner
  - avoiding excessive vectoring or over-controlling
Comments:
	1	2	3	4	5	6	7	8

8. Correcting Own Errors in a Timely Manner **Comments:**	1	2	3	4	5	6	7	8
9. Overall Attention and Situation Awareness Scale Rating **Comments:**	1	2	3	4	5	6	7	8

Prioritizing

	1 2 3 4 5 6 7 8
10. Taking Actions in an Appropriate Order of Importance - resolving situations that need immediate attention before handling low priority tasks - issuing control instructions in a prioritized, structured, and timely manner **Comments:** _____ _____ _____	1 2 3 4 5 6 7 8
11. Preplanning Control Actions - scanning adjacent sectors to plan for inbound traffic - studying pending flight strips in bay **Comments:** _____ _____ _____	1 2 3 4 5 6 7 8
12. Handling Control Tasks for Several Aircraft - shifting control tasks between several aircraft when necessary - avoiding delays in communications while thinking or planning control actions **Comments:** _____ _____ _____	1 2 3 4 5 6 7 8
13. Marking Flight Strips While Performing Other Tasks - marking flight strips accurately while talking or performing other tasks - keeping flight strips current **Comments:** _____ _____ _____	1 2 3 4 5 6 7 8
14. Overall Prioritizing Scale Rating **Comments:** _____ _____ _____	1 2 3 4 5 6 7 8

Providing Control Information

	1 2 3 4 5 6 7 8
15. Providing Essential Air Traffic Control Information - providing mandatory services and advisories to pilots in a timely manner - exchanging essential information **Comments:** _____ _____ _____	1 2 3 4 5 6 7 8
16. Providing Additional Air Traffic Control Information - providing additional services when workload is not a factor - exchanging additional information **Comments:** _____ _____ _____	1 2 3 4 5 6 7 8
17. Overall Providing Control Information Scale Rating **Comments:** _____ _____ _____	1 2 3 4 5 6 7 8

Technical Knowledge

18. Showing Knowledge of LOAs and SOPs
   - controlling traffic as depicted in current LOAs and SOPs
   - performing hand-off procedures correctly
Comments:
	1 2 3 4 5 6 7 8

19. Showing Knowledge of Aircraft Capabilities and Limitations
   - avoiding clearances that are beyond aircraft performance parameters
   - recognizing the need for speed restrictions and wake turbulence separation
Comments:
	1 2 3 4 5 6 7 8

20. Overall Technical Knowledge Scale Rating **Comments:**	1 2 3 4 5 6 7 8

Communicating

21. Using Proper Phraseology
   - using words and phrases specified in ATP 7110.65
   - using ATP phraseology that is appropriate for the situation
   - avoiding the use of excessive verbiage
Comments:
	1 2 3 4 5 6 7 8

22. Communicating Clearly and Efficiently
   - speaking at the proper volume and rate for pilots to understand
   - speaking fluently while scanning or performing other tasks
   - clearance delivery is complete, correct, and timely
   - providing complete information in each clearance
Comments:
	1 2 3 4 5 6 7 8

23. Listening for Pilot Readbacks and Requests
   - correcting pilot readback errors
   - processing requests correctly in a timely manner
Comments:
	1 2 3 4 5 6 7 8

24. Overall Communicating Scale Rating **Comments:**	1 2 3 4 5 6 7 8

Instructions for questions 25-35

The following questions have a scale ranging from 1 to 10, where 1 represents "extremely low," "extremely infrequent," "strongly disagree," etc., and 10 represents the other extreme of the spectrum.

These questions are the same as we have asked the controller after the scenario. We would like you to give us your impression of how these questions will be rated by the controller.

25. Please circle the number that best describes the controller's **preference for vertical separation.** **Comments:**	no vertical separation	1 2 3 4 5 6 7 8 9 10	always vertical separation

26. Please circle the number that best describes the controller's **preference for separation through "vectoring."** **Comments:**	no vector separation	1 2 3 4 5 6 7 8 9 10	always vector separation

27. Please circle the number that best describes the controller's **preference for speed control.** **Comments:**	no speed control	1 2 3 4 5 6 7 8 9 10	always speed control

28. Please circle the number below that best describes **how hard the controller was working** during this scenario. **Comments:**	not hard	1 2 3 4 5 6 7 8 9 10	extremely hard

29. Please circle the number that best describes **how well the controller controlled traffic** during this scenario. **Comments:**	extremely poor	1 2 3 4 5 6 7 8 9 10	extremely well

NASA TLX

30. Please circle the number that best describes the **mental demand** during this scenario.	extremely low	1 2 3 4 5 6 7 8 9 10	extremely high
31. Please circle the number that best describes the **physical demand** during this scenario.	extremely low	1 2 3 4 5 6 7 8 9 10	extremely high
32. Please circle the number that best describes the **temporal demand** during this scenario.	extremely low	1 2 3 4 5 6 7 8 9 10	extremely high
33. Please circle the number that best describes the overall **performance** during this scenario.	extremely low	1 2 3 4 5 6 7 8 9 10	extremely high
34. Please circle the number that best describes the **effort** during this scenario.	extremely low	1 2 3 4 5 6 7 8 9 10	extremely high
35. Please circle the number that best describes the level of **frustration** during this scenario.	extremely low	1 2 3 4 5 6 7 8 9 10	extremely high

Appendix B
Situation Awareness Global Assessment Technique Queries

(on the provided sector map)
1. Enter the location of all aircraft
        aircraft in track control
        other aircraft in sector
        aircraft that will be in track control in next 2 minutes

2. Enter aircraft callsign (for aircraft highlighted of those entered on sector map in query 1)

3. Enter aircraft altitude (for aircraft highlighted of those entered on sector map in query 1)

4. Enter aircraft groundspeed (for aircraft highlighted of those entered on sector map in query 1)

5. Enter aircraft heading (for aircraft highlighted of those entered on sector map in query 1)

6. Enter aircraft's next sector (for aircraft highlighted of those entered on sector map in query 1)
        02   49   67
        15   57
        16   58
        35   65
        landing in sector

7. Enter aircraft's current direction of change in each column (for aircraft highlighted of those entered on sector map in query 1)

Altitude change	Turn
climbing	right turn
descending	left turn
level	straight

8. Enter the aircraft type (for aircraft highlighted of those entered on sector map in query 1)

9. Enter the aircraft's activity in sector (for aircraft highlighted of those entered on sector map in query 1)
                enroute
                inbound to airport
                outbound from airport

10. Which pairs of aircraft have lost or will currently lose separation if they stay on their current (assigned) courses?

11. Which aircraft have been issued clearances that have not been completed?

12. Did the aircraft receive its clearance correctly? (for each of those entered in query 11)

13. Which aircraft are currently conforming to their clearances? (for each of those entered in query 11)

14. Which aircraft will be handed off to another sector/facility in the next 2 minutes?

15. Enter the aircraft which are not in communication with you.

16. Enter the aircraft that will violate special airspace separation standards if they stay on their current assigned paths.

17. Which aircraft is weather currently an impact on or will be an impact on in the next 5 minutes?

Appendix C
NASA-TLX

Mental Demand

How much mental and perceptual activity is required (e.g., thinking, deciding, calculating, remembering, looking, searching)? Is the task easy or demanding, simple or complex, or exacting or forgiving?

Low |————————————————————————————| High

Physical Demand

How much physical activity is required (e.g., pushing, turning, controlling, activating)? Is the task easy or demanding, slow or brisk, slack or strenuous, or restful or laborious?

Low |————————————————————————————| High

Temporal Demand

How much time pressure do you feel due to the rate or pace at which the tasks or task elements occurred? Is the pace slow and leisurely or rapid and frantic?

Low |————————————————————————————| High

Performance

How successful do you think you are in accomplishing the goals of the task? How satisfied are you with your performance in accomplishing these goals?

Good |————————————————————————————| Poor

Effort

How hard did you have to work (mentally and physically) to accomplish this level of performance?

Low |————————————————————————————| High

Frustration

How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed, and complacent do you feel in performing the task?

Low |————————————————————————————| High

Appendix D
Post-Scenario Questionnaire

1. Please circle the number that best describes **how realistic the simulation** was.

    extremely unrealistic   1 2 3 4 5 6 7 8 9 10   extremely realistic

    Comments: _____
    _____
    _____
    _____

2. Please circle the number that best describes if the **ATWIT device interfered** with controlling traffic.

    no interference   1 2 3 4 5 6 7 8 9 10   extreme interference

    Comments: _____
    _____
    _____
    _____

3. Please circle the number that best describes **how well the simulation-pilots responded** to your clearances in terms of traffic movement and call-backs.

    extremely poor   1 2 3 4 5 6 7 8 9 10   extremely well

    Comments: _____
    _____
    _____
    _____

4. Do you have any other comments about your experiences during the simulation?

    Comments: _____
    _____
    _____
    _____

5. Please circle the number below that best describes **how hard you were working** during this scenario.

    not hard   1 2 3 4 5 6 7 8 9 10   extremely hard

    Comments: _____
    _____
    _____
    _____

6. Please circle the number that best describes **how well you controlled traffic** during this scenario.

    extremely poor   1 2 3 4 5 6 7 8 9 10   extremely well

    Comments: _____
    _____
    _____
    _____

7. Please circle the number that best describes **how difficult** this scenario was.

extremely easy   1 2 3 4 5 6 7 8 9 10   extremely difficult

Comments: _____

_____

_____

_____

---

1. Please circle the number that best describes **how realistic the simulations** were.

extremely unrealistic   1 2 3 4 5 6 7 8 9 10   extremely realistic

Comments: _____

_____

_____

_____

---

2. Please circle the number that best describes if the **ATWIT device interfered** with controlling traffic.

no interference   1 2 3 4 5 6 7 8 9 10   extreme interference

Comments: _____

_____

_____

_____

---

3. Please circle the number that best describes **how well the simulation-pilots responded** to your clearances in terms of traffic movement and call-backs.

extremely poor   1 2 3 4 5 6 7 8 9 10   extremely well

Comments: _____

_____

_____

_____

---

4. Please circle the number that best describes if the **hands-on training for each scenario was adequate**.

not adequate   1 2 3 4 5 6 7 8 9 10   adequate

Comments: _____

_____

_____

_____

---

5. Is there anything about the study that we should have asked or that you would like to comment about?

Comments: _____

_____

_____

_____

Appendix E
SME SAGAT Data Evaluation Form

Subject _____   **Condition** _____   **Scenario** _____ **Trial** _____   **Stop number** _____

* if aircraft stays on current (assigned) path

	Aircraft	Track Control	Vertical velocity	Turning	Next Sector	Sector airspace violation in next 2 min*	Not in comm with sector	Will violate SUA next 2 min*	Weather will impact in next 5 min*
1		my control in next 2 min other in sector	level climbing descending	straight left right					
2		my control in next 2 min other in sector	level climbing descending	straight left right					
3		my control in next 2 min other in sector	level climbing descending	straight left right					
4		my control in next 2 min other in sector	level climbing descending	straight left right					
5		my control in next 2 min other in sector	level climbing descending	straight left right					
6		my control in next 2 min other in sector	level climbing descending	straight left right					
7		my control in next 2 min other in sector	level climbing descending	straight left right					
8		my control in next 2 min other in sector	level climbing descending	straight left right					

Which pairs of aircraft have lost or will lose separation in the next 2 minutes if they stay on their current (assigned) courses?

Which aircraft have assignments

(clearances that are not yet complete?)

	Received correctly?	Conforming to assigned clearance?
_____	Y / N	Y / N
_____	Y / N	Y / N
_____	Y / N	Y / N
_____	Y / N	Y / N
_____	Y / N	Y / N

E-1

Appendix F
Exit Questionnaire

1.	Please circle the number that best describes **how realistic the simulations** were. **Comments:**	extremely unrealistic	1 2 3 4 5 6 7 8 9 10 extremely realistic

2.	Please circle the number that best describes if the **ATWIT device interfered** with controlling traffic. **Comments:**	no interference	1 2 3 4 5 6 7 8 9 10 extreme interference

3.	Please circle the number that best describes **how well the simulation-pilots responded** to your clearances in terms of traffic movement and call-backs. **Comments:**	extremely poor	1 2 3 4 5 6 7 8 9 10 extremely well

4.	Please circle the number that best describes if the **hands-on training for each scenario was adequate**. **Comments:**	not adequate	1 2 3 4 5 6 7 8 9 10 adequate

5.	Is there anything about the study that we should have asked or that you would like to comment about? **Comments:**